Hybrid Correlation Models For Bond Breaking Based On Active Space Partitioning

A Thesis
Presented to
The Academic Faculty

by

Artem D. Bochevarov

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

School of Chemistry and Biochemistry
Georgia Institute of Technology
August 2006

Hybrid Correlation Models For Bond Breaking Based On Active Space Partitioning

Approved by:

Professor C. David Sherrill, Adviser
(School of Chemistry & Biochemistry)
Georgia Institute of Technology

Professor Rigoberto Hernandez
(School of Chemistry & Biochemistry)
Georgia Institute of Technology

Professor Jean-Luc Brédas
(School of Chemistry & Biochemistry)
Georgia Institute of Technology

Professor Thomas Orlando
(School of Chemistry & Biochemistry)
Georgia Institute of Technology

Professor Mei-Yin Chou
(School of Physics)
Georgia Institute of Technology

Date Approved: June 1, 2006

ACKNOWLEDGEMENTS

The time spent in the graduate school at Georgia Institute of Technology was the period of deep self-introspection and extensive learning. I was influenced, taught and helped by many people, whose names are given below in the alphabetical order. To all of them I express my sincere gratitude.

Micah L. Abrams, William H. Adams, Roye Albrighton, Jon Anderson, Marc C. Baker, Tony Banks, Hugh Banton, Richard Barbieri, Samuel Beckett, Adrian Belew, Arnold Bennett, Loïc Bernardeau, Pascal Bertrand, Klaus Blasquiz, the Bochevarovs, Mick Brockett, Bill Bruford, Edward Bulwer-Lytton, Pietr Luigi Calderoni, Phil Collins, Joseph Conrad, Sebastian Constard, Pierre Corneille, T. Daniel Crawford, Ernie Croot, Renato D'Angelo, Michi Dei Rossi, Stanislav Derevyanko, Charles Dickens, Francesco DiGiacomo, R. M. W. Dixon, David Donson, Fyodor Dostoyevsky, Colin Edwin, Keith Emerson, Jonas Engdegård, Francesco Evangelista, Guy Evans, Henry Fielding, Fish, Allan Freeman, Guillaume Fountaine, John Fowles, Robert Fripp, Anthony Gabard, Peter Gabriel, Benoît Gaignon, René Garber, David Gilmour, George Gissing, Brian Greene, Lionel B. Guichard, Steve Hackett, Peter Hammill, Thomas Hardy, Gavin Harrison, So Hirata, Johan Högberg, Anna Holmgren, Ron Howden, Steve Howe, Jean Baptiste Itier, David Jackson, Derek Jakobi, Jonas Jarvholm, Thomas Johnson, Garry Kasparov, Mark Kelly, Olga Kittlerová, Werner Kutzelnigg, Greg Lake, Teddy Lasry, Tony Levin, Tord Lindman, Mery Looze, Jean Pierre Louveton, David Lucas, Anatoly Luzanov, Jean-Luc Manderlier, Nick Mason, W. Somerset Maugham, Guy de Maupassant, Sergey Maximov, George Meredith, Arthur Miller, Derek Moore, Ian Mosley, Vladimir Nabokov, Gianni Nocenzi, Vittorio Nocenzi, Claude Olmos, Jeppe Olsen, Mattias Olsson, Amy Overmyer, Toni Pagliuca, Joseph Paldus, Carl Palmer,

Alexander Popov, Yaroslav Prilepsky, Charles Reade, Samuel Richardson, Steve Rothery, Mike Rutherford, Seth Schneer, John S. Sears, Alex Semenyaka, Anastasia Senenko, William Shakespeare, Manu Sharma, Richard B. Sheridan, Derek Sherinian, C. David Sherrill, Sergey Shipov, Tobias Smollett, Kyril Solntsev, Chris Squire, Laurence Sterne, Aldo Tagliapietra, Berhane Temelso, William M. Thackeray, Yar Tikhiy, Marcello Todaro, Lev Tolstoy, Jannik Top, Pete Trewavas, Edward F. Valeev, Christian Vander, Voltaire, Edward Wagencknecht, Rick Wakeman, Roger Waters, Ryan West, John Wetton, Oscar Wilde, Steven Wilson, Richard Wright, Andrey Zaitsevsky.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The understanding and proper characterization of bond-breaking processes are vital to many areas of chemistry. Indeed, chemical processes are often defined as those in which chemical bonds are made or broken. It is unfortunate then that even the qualitatively correct theoretical description of the dissociation process is difficult to achieve within the scope of all but the most sophisticated quantum-chemical methods.

The work presented in this thesis is dedicated to developing inexpensive quantum-chemical models that are able to produce smooth and physically correct potential energy curves for the dissociation of single covalent bonds. It is well known that the energies produced by many *ab initio* theories scaling as the fifth order with the system size (for instance, second-order Møller-Plesset (MP2) and Epstein-Nesbet perturbation theories) diverge at large interatomic separations. We show that the divergent behavior of such perturbation schemes is due to a small number of terms in the energy expressions. Then, we demonstrate that the self-consistent replacement of these terms by their analogs from the coupled cluster theory (such as CCSD) allows one to redress the erroneous behavior of the perturbation theories without the damage to the overall scaling.

We also investigate the accuracy of these *hybrid* perturbation theory-coupled cluster theories near equilibrium geometry. Judging from the computed spectroscopic constants and shapes of the potential energy curves, one such model, denoted MP2-CCSD(II) in this work, performs consistently better than the MP2 theory at essentially the same computational cost.

CHAPTER I

INTRODUCTION

1.1 Development of new theoretical tools in quantum chemistry

Quantum chemistry is a science that has sprung from a more general physical discipline, quantum mechanics, and is traditionally occupied with the explanation of electronic structure of molecules by means of quantum-mechanical models. Such models, aimed at molecules and their interactions, are usually characterized as quantum-chemical. Computational chemistry, a field of study closely related to quantum chemistry, deals with the efficient computer implementation of existent quantum-chemical models and their application for the solution of specific chemical problems. This terminology is not universally agreed upon, but in this thesis we shall adhere to the definitions made above. The distinction between the quantum and the computational directions in chemical research is often helpful when it comes to the choice of methodology and interpretation of the results. Quantum chemistry adopts a *physical* approach, striving for the understanding of a phenomenon and its qualitative explanation in terms of a physically sound mathematical model. Computational chemistry is more concerned with obtaining an accurate *numerical* result and then using it to explain the studied phenomenon or property. Some authors [1, 2] use the term 'computational quantum chemistry' apparently trying to refer to both approaches or as an embellishment of the term 'computational chemistry', but in our opinion its meaning is somewhat confusing. Perhaps, this is one of the reasons why the Center for Computational Quantum Chemistry (CCQC) at the University of Georgia was recently renamed into Center for Computational Chemistry (CCC) which unambiguously marks off its research area.

This thesis combines and describes several projects dedicated to the construction of

new theoretical methods and algorithms within the framework of quantum chemistry. In our works we lay a particular emphasis on innovation and development of novel approaches rather than the application of well-tested theories to new molecular systems.

1.2 Thesis structure

Chapter 2 gives the key points on basic quantum chemistry and serves as a general introduction. Although this material is available in many introductory texts (see, for example, Refs. [1, 3, 4, 5]) its appearance here would be justified even if it were only to specify our notation used in the rest of the thesis. Each subsequent chapter presents a separate project preceded by a brief introduction. Every such project was either published as an original scientific paper or (at the moment of writing this thesis) submitted to publication. Therefore a lot of material presented in chapters 3-7 intersects with that in the corresponding publications but occasionally contains some additional unpublished data or discussions.

Chapters 3 and 4 deal with the work which gives the title to the thesis: the construction of inexpensive hybrid quantum chemical models which may be used as an alternative to the standard methods for bond-breaking. Chapter 5 describes our work on the so-called nuclear orbitals, the one-particle functions analogous to the molecular electronic orbitals, and the non-adiabatic functions which include these nuclear orbitals. Chapter 6 delineates a diagrammatic algorithm designed to handle the second-quantized expressions and to facilitate the work with some symbolic expressions in quantum chemistry. Chapter 7 presents an investigation of some interesting mathematical constructions which follow trivially from the Löwdin's partitioning technique.

CHAPTER II

THE BASIC DEFINITIONS AND NOTIONS OF QUANTUM CHEMISTRY

2.1 The main techniques of quantum chemistry

The main objective of quantum chemistry is to solve the electronic Schrödinger equation

$$\hat{H}|\Psi_k\rangle = E_k|\Psi_k\rangle \tag{1}$$

which is written in the Hilbert space for the many-electron Hamiltonian $\hat{H}$ often containing only electrostatic interactions and thus having the following form in atomic units

$$\hat{H} = -\frac{1}{2}\sum_i \nabla_i^2 - \sum_{i,A}\frac{Z_A}{r_{iA}} + \sum_{A>B}\frac{Z_A Z_B}{R_{AB}} + \sum_{i>j}\frac{1}{r_{ij}}. \tag{2}$$

The letters i, j and A, B refer to the electron and nuclear coordinates, respectively. As the analytical solution of (1-2) is possible only for one-electron system, quantum chemistry tries to find the best approximation to $|\Psi_k\rangle$ and/or E_k. The so-called *ab initio* approach does not approximate the Hamiltonian (2) or its matrix elements by some simple expressions that typically involve *ad hoc* parameters, which is a course taken by the *semi-empirical* quantum chemistry. The popular density functional theory (DFT) methods attempt to modify the Hamiltonian so that the electron density computed with the simplest possible form of the wave function coincides with the exact electron density. In this thesis, only *ab initio* models will be considered. When the wave function for the ground state is sought (then $k = 0$), it is often given some functional form. Then the approximate function $|\Psi_0'\rangle$ which minimizes some energy functional

$$E = \mathcal{F}\{|\Psi_0'\rangle\} \tag{3}$$

is considered the best approximation to $|\Psi_0\rangle$ within this functional form. The most common energy functional (3) is due to Rayleigh and Ritz:

$$E = \langle\Psi_0'|H|\Psi_0'\rangle / \langle\Psi_0'|\Psi_0'\rangle. \tag{4}$$

The simplest representation of the N-electron trial wave function $|\Phi\rangle$ is a determinant

$$|\Phi(1,2,...,N)\rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_1(r_1) & \varphi_2(r_1) & ... & \varphi_N(r_1) \\ \varphi_1(r_2) & \varphi_2(r_2) & ... & \varphi_N(r_2) \\ ... & ... & ... & ... \\ \varphi_1(r_N) & \varphi_2(r_N) & ... & \varphi_N(r_N) \end{vmatrix} \tag{5}$$

constructed from one-electron functions $\varphi_q(r)$ called orbitals. Typically, a linear combination of such determinants

$$|\Psi\rangle = \sum_k c_k |\Phi_k\rangle \tag{6}$$

is used, in which the variational parameters are either the orbitals $\varphi_q(r)$ or the coefficients c_k, or both. The determinants $|\Phi_k\rangle$ should differ in at least one orbital $\varphi_q(r)$ in order to be linearly independent. The methods that vary some functional are called *variational*. Such are the popular Hartree-Fock (HF) [6, 7, 8, 9, 10] and the configuration interaction (CI) [11, 12] methods.

There is a wide group of other methods that do not perform any functional minimization. Such *non-variational* methods usually substitute $|\Psi_k\rangle$ in the Schrödinger equation (1) by some approximation $|\Psi_k'\rangle$ containing parameters. The best approximation is found for the parameters that satisfy the series of *projected* Schrödinger equations

$$\langle\Xi_l|\hat{H}|\Psi_k'\rangle = E_k\langle\Xi_l|\Psi_k'\rangle \tag{7}$$

where $\{\langle\Xi_l|\}$ is some set of linearly independent vectors in the Hilbert space. This is the approach adopted, for example, by the coupled-cluster (CC) methods [13, 14, 15, 16, 17, 18] and the equation-of-motion coupled cluster (EOM-CC) methods [19, 20, 21, 22, 23, 24].

Numerous types of perturbation theory [25, 26, 27, 28] used in quantum chemistry assume that the looked-for wave function $|\Psi\rangle$ and the energy E can be decomposed in some small parameter ε:

$$|\Psi\rangle = |\Psi^{(0)}\rangle + \varepsilon|\Psi^{(1)}\rangle + \varepsilon^2|\Psi^{(2)}\rangle + ..., \tag{8}$$

$$E = E^{(0)} + \varepsilon E^{(1)} + \varepsilon^2 E^{(2)} + ... \tag{9}$$

Often the equations 'decouple' and are solved independently for each order of smallness. Various forms of perturbation (characterized by ε) and the ways to express it lead to different forms of the perturbation expansions.

2.2 The fundamental difficulties of quantum chemistry

The initial obstacles in quantum chemistry emerge long before we arrive at the Schrödinger equation (1). In writing (1), we had to make two drastic approximations: first, we removed all relativistic terms from the Hamiltonian and consequently, we treat $|\Psi^{(0)}\rangle$ as a one-component vector and not a spinor. Second, we applied the Born-Oppenheimer approximation [29] which allows us to decouple the electronic motion from the nuclear one. Although for the molecules with the light atoms these approximations typically lead to small errors in numerical results, in general the simple equation (1) has to be corrected for the relativistic and non-adiabatic effects. Numerous such corrections or even first-principle treatments exist, but except in Chapter 5, where we deal with the non-Bon-Oppenheimer effects, we will concern ourselves only with the approximation to the solutions of the simple equation (1).

Even though such approximations simplify greatly the task of finding the electronic structure of molecules, solving the Schrödinger equation (1) is still not easy. The main reason is that for the effective treatment of the electron correlation (defined as energy 'not captured' by the HF method) one needs to employ a large number of determinants (often tens and hundreds of millions or more) in (6). The total number of determinants one can

construct from a given set of orbitals often scales factorially with the number of these orbitals [30], and soon reaches the limit which makes the calculations prohibitively expensive to carry out. Much of quantum chemistry nowadays is concerned with how to select only most relevant determinants without greatly compromising the accuracy of the calculation. But even if we have chosen the right set of the determinants at the equilibrium geometry, it may be seriously incomplete at large interatomic separations, for instance, when two or more electronic states are close in energy.

Another important problem is the complexity of quantum-chemical expressions which result when one derives expressions involving an arbitrary set of determinants of arbitrary structure. Such expressions are often even difficult to write down on paper since they may be too long. Decades have passed since the introduction of expansion (6) before quantum chemists found ways of efficient computer implementation which bypass the explicit formulation of such methods as CI and CC.

In the next chapter we will present the Hartree-Fock method which is a necessary step in the implementation of almost any ab initio procedure.

2.3 The Hartree-Fock method

The significance of the Hartree-Fock method is due to its ability to find reasonably accurate one-electron functions (orbitals) which may be used in the construction of many-electron wave functions. Our derivation here is adopted after an ingenious method of Adams [31] which is an entirely operator approach.

In order we embark on this course, we must introduce the one-electron density operator, sometimes called density matrix

$$\rho(1;1') = \int dx_2 dx_3 ... dx_N |\Psi(1,2,3...,N)\rangle\langle\Psi(1',2,3...,N)| \tag{10}$$

which allows one to write an expectation value $\bar{A}$ of any one-particle operator $\hat{A}(1)$:

$$\bar{A} = \int dx_1 \rho(1;1)\hat{A}(1). \tag{11}$$

This is often written as

$$\bar{A} = \mathrm{Sp}_{(1)} \left[\rho(1;1) \hat{A}(1) \right]. \tag{12}$$

The operation $\mathrm{Sp}_{(\mathrm{p..q})}$ (from German Spur) performs a contraction (integration or summation, depending on context) in variables from p to q.

Suppose that the one-particle density operator of the N-electron system is defined by the following simple formula

$$\rho(1;1') = \sum_{k}^{N} |\varphi_k(1)\rangle\langle\varphi_k(1')|, \tag{13}$$

with the the idempotency property imposed on it

$$\rho(1;1')\rho(1';1'') = \rho(1;1''). \tag{14}$$

Suppose also that the two-particle density matrix $\rho(1,2;1',2')$ is defined as

$$\rho(1,2;1',2') = \rho(1;1')\rho(2;2') - \rho(1;2')\rho(2;1'). \tag{15}$$

These definitions of the one- and two-particle density matrices are consistent with the wave-function derivation of the Hartree-Fock method, since these density matrices follow from the Slater determinant, which is the HF wave function. Then the electronic energy of the HF method will be written as

$$E = \frac{1}{2}\mathrm{Sp}_{(1)}\rho(1;1)(2h(1) + J(1) - K(1)). \tag{16}$$

The 'diagonal' density operator $\rho(1;1)$ is often written simply as $\rho(1)$. In (16) $h(1)$ is the one-electron part of the Hamiltonian defined by

$$h(1) = \frac{1}{2}\nabla_1^2 - \sum_A \frac{Z_A}{r_{1A}}, \tag{17}$$

whereas $J(1)$ and $K(1)$ are the summed analogs of the well-known Coulomb and exchange operators, respectively:

$$J(1) = \sum_{k}^{N} J_k(1) = \int dx_2 \; \rho(2)/r_{12}, \tag{18}$$

$$K(1) = \sum_k^N K_k(1) = \int dx_2 \; \rho(1;2)/r_{12}\mathcal{P}_{12}. \tag{19}$$

$\mathcal{P}_{12}$ in the last expression is the permutation operator which changes coordinate 1 into coordinate 2 in any ket-vector it acts upon. For example,

$$\mathcal{P}_{12}\rho(1) = \rho(2;1), \tag{20}$$

and consequently,

$$K(1)\rho(1) = \sum_k^N K_k(1) = \int dx_2 \; \rho(1;2)\rho(2;1)/r_{12}. \tag{21}$$

Let us now determine which ρ makes the energy (16) stationary with respect to the perturbation λ. The perturbed density operator will have the form

$$\rho' = \rho + \lambda\rho^{(1)} + \lambda^2\rho^{(2)} + ... \tag{22}$$

The idempotency condition (14) leads us to

$$\rho\rho^{(1)} + \rho^{(1)}\rho = \rho^{(1)}, \tag{23}$$

$$\rho\rho^{(2)} + \rho^{(1)}\rho^{(1)} + \rho^{(2)}\rho = \rho^{(2)}. \tag{24}$$

Now we introduce the operator

$$v = (1-\rho)\Delta\rho \tag{25}$$

where Δ is a small arbitrary operator. Then it is not difficult to prove that $v + v^\dagger$ equals $\rho^{(1)}$. Indeed, inserting $v + v^\dagger$ in the left hand side of equation (23), we obtain after simple algebraic operations $v + v^\dagger$ on the left hand side. One may also ascertain, using the similar operations with equation (24), that $vv^\dagger - v^\dagger v + v' + v'^\dagger$ equals $\rho^{(2)}$. Here v' is some other small arbitrary operator: $v' = (1-\rho)\Delta'\rho$.

The energy stationarity requirement is

$$E^{(1)} = \left(\frac{\partial E}{\partial\lambda}\right)\Bigg|_{\lambda=0} = 0. \tag{26}$$

Substituting (22) into (16) and making use of the obvious equality

$$\mathrm{Sp}\left[\rho^{(0)}(J^{(1)}-K^{(1)})\right]=\mathrm{Sp}\left[\rho^{(1)}(J-K)\right] \tag{27}$$

where the coordinate indices are not explicitly shown for not to clutter the formula and introducing the Fock operator F

$$F(1)=h(1)+J(1)-K(1) \tag{28}$$

we derive the stationarity condition in terms of Δ:

$$E^{(1)}=\mathrm{Sp}\left[(h+J-K)\rho^{(1)}\right]=\mathrm{Sp}\left[F(v+v^{\dagger})\right]=\mathrm{Sp}\left[\rho F(1-\rho)\Delta+(1-\rho)F\rho\Delta^{\dagger})\right]=0. \tag{29}$$

For arbitrary Δ

$$\rho F(1-\rho)=0 \tag{30}$$

will hold. This is true only if

$$\rho F=F\rho \tag{31}$$

which is the sought-for Hartree-Fock equation. It is more often written in terms of the molecular orbitals (MOs) φ_k:

$$F\varphi_k=\varepsilon_k\varphi_k \tag{32}$$

which follows from (31) if we recall that the commuting operators have the same set of eigenfunctions. This operator method also allows an elegant investigation of the second derivative of energy E with respect to perturbation λ which leads to the so-called orbital stability conditions [31, 32].

A posteriori, although usually it is done *a priori*, one can observe that the simple expressions for the density matrices (17), (15) are possible only if the wave function of the Hartree-Fock method is a determinant constructed from the orbitals $\varphi_q(r)$. Thus, the Hartree-Fock method finds the set of (occupied) orbitals which minimizes the Rayleigh-Ritz functional (4).

Each orbital ε_k certainly depends on the spatial and the spin coordinates of the electron it describes. It is in fact a product of the spatial part and the spin part, the latter represented by only two functions, commonly known as α and β. In many cases the spatial function for the α- and β-electrons may be the same. When this is the case for the closed shell, such approach is called the restricted Hartree-Fock (RHF) method. When it is an open shell, the method is called open shell restricted Hartree-Fock (ROHF). When the spatial functions of the orbitals are not restricted to be the same for the α- and β-electrons, the method is called the unrestricted Hartree-Fock (UHF).

The solution of integro-differential equations (32), and consequently, the representation of the orbitals ε_k is conveniently done in a discrete basis set. For all the calculations presented in this thesis we employed basis sets of the Gaussian type. The structure and the nomenclature of the quantum-chemical basis sets are rather complicated and full of details and exceptions, and therefore we refer the reader to the special literature (**??**).

2.4 The electron correlation

Since the Hartree-Fock theory is effectively a one-electron theory it describes each electron in the average electric field created by other electrons and the nuclei. More sophisticated methods take into account *electron correlation*, that is, when electrons are 'aware' of the immediate positions of one another. Such methods produce lower and more accurate energy of the system and the difference between the energy of this method and the Hartree-Fock energy is called the *correlation energy*.

The electron correlation methods must necessarily abandon the simple picture of each electron described by its own orbital. However, the MO theory is so attractive in its simplicity that many electron correlation methods make heavy use of it. To make the connection from the Hartree-Fock theory to the exact solution of the Schrödinger equation we notice that the linear combination of all possible determinants (6) constructed with the complete set of the MOs can represent the exact wave function granted the appropriate coefficients

c_k are found.

There is a great variety of electron correlation methods, but in this introduction we can only succinctly review those which are directly connected with the material given in the following chapters.

The configuration interaction method seeks the wave function in the form (6). In theory, the number of determinants should be infinite. In practice, however, we find that a modest set of determinants can already give a significant improvement over the HF theory. In the simplest realization of the CI method, its equations are obtained by the projection of the Schrödinger equation with (6) as a wave function by N determinants which are included in the linear combination:

$$\langle \Phi_i | H | \Phi_j \rangle c_{ij} = E\delta_{ij}, \qquad i,j \in 1...N. \tag{33}$$

This set of linear equations in c_{ij} is obviously a secular problem which is easily solvable for N not exceeding a few thousand. Unfortunately, the required number of determinants may count in millions, and therefore other, indirect, but more efficient methods are used instead.

The coupled cluster method is similar to CI in that it approximates the wave function as a linear combination of determinants. However, the coefficients of this linear combination may be products. Such a decomposition no longer allows for a linear system of equations, but it makes the approximation of the exact function more efficient: fewer coefficients need to be found than in CI for the comparable accuracy. The coupled cluster wave function is written as an *exponential ansatz*:

$$|\Psi\rangle = e^{\hat{T}} |\Phi_0\rangle \tag{34}$$

where $|\Phi_0\rangle$ is the Hartree-Fock wave function and T has the form

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + ... + \hat{T}_N. \tag{35}$$

The $\hat{T}_k$ operator replaces k occupied orbitals in the Hartree-Fock determinant with k virtual ones and adds the resulting determinants $|\Phi(k)_i\rangle$ together with some linear coefficients

$\{t(k)\}$:

$$\hat{T}(k)|\Phi_0\rangle = \sum_i t(k)_i |\Phi(k)_i\rangle. \tag{36}$$

The unknown coefficients $t(k)_i$ may be found from the set of projected Schrödinger equations (7), with (34) as the wave function.

The third very popular electron correlation method is the Møller-Plesset perturbation theory. It uses decompositions (8) and (9) starting from the Hartree-Fock determinant and energy as the zero-order approximation. The first perturbative correction to the ground state energy is zero, and the second one is defined by the formula:

$$E_{\rm corr}^{\rm MP2} = \frac{1}{4}\sum_{ijab} \frac{|(ij|ab) - (ib|aj)|^2}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b}, \tag{37}$$

where $\langle ij|ab\rangle$ is the so called two-electron integral in the chemical notation

$$\langle ij|ab\rangle = \int dx_1 \int dx_2\, \varphi_i(x_1)\varphi_j(x_1)(r_{12}^{-1})\varphi_a(x_2)\varphi_b(x_2) \tag{38}$$

and ε_i is the orbital energy defined in (32). The indices i and j denote the occupied orbitals whereas a and b denote virtual orbitals. The first correction to the wave function is not zero:

$$|\Psi^{(1)}\rangle = \frac{1}{4}\sum_{ijab} \frac{(ij|ab) - (ib|aj)}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b} |\Phi_{ij}^{ab}\rangle. \tag{39}$$

Here $|\Phi_{ij}^{ab}\rangle$ is the twice excited determinant obtained from the ground state determinant by substituting the orbitals φ_i and φ_j by the virtual orbitals φ_a and φ_b.

2.5 The second quantization and the diagrammatic technique

The method of second quantization is very general and is used extensively in many areas of physics. In this introduction we will describe how it is applied in quantum chemistry. Detailed explanations of the second quantization and the diagrammatic technique (which is another level of abstraction on top of second quantization and will be explained in the following chapter) may be found in several excellent introductory texts [33, 34, 35, 36, 37].

The discussion of second quantization in quantum chemistry may be commenced by the observation that it is not at all necessary to work with determinants and other wave functions. First, we can stay in the so-called first quantization and deal with traditional operators only. Instead of using wave functions we can work with the corresponding density matrices that contain all the information about our system. Another approach, called the second quantization, obliterates the use of wave functions in a different way. The analogs of wave functions here are special strings of *creation* and *annihilation* operators. More than that, each operator from the first quantization may also be expressed through the creation and annihilation operators. The resulting space of operators is referred to as the *Fock space*.

In order to relate the first quantization to the second one, we define the linear creation operator $a_i^\dagger$ as an operator that creates a particle (an electron) which is described by the orbital φ_i:

$$a_i^\dagger|0\rangle = |\varphi_i\rangle. \tag{40}$$

The vector $|0\rangle$ is the so-called vacuum vector, a determinant of size zero, and $|\varphi_i\rangle$ is a determinant of size one. More generally, $a_i^\dagger$ acts on a determinant of arbitrary size increasing its size by one and giving the new orbital index 1:

$$a_i^\dagger \frac{1}{\sqrt{N!}}\det|\varphi_a(1)\varphi_b(2)...\varphi_q(N)| = \frac{1}{\sqrt{(N+1)!}}\det|\varphi_i(1)\varphi_a(2)\varphi_b(3)...\varphi_q(N+1)|. \tag{41}$$

The Hermitian conjugate of $a_i^\dagger$ is the annihilation operator a_i which destroys the orbital φ_i in the determinant it acts upon:

$$a_i \frac{1}{\sqrt{N!}}\det|\varphi_a(1)\varphi_b(2)...\varphi_i(L)...\varphi_q(N)| = \frac{(-1)^P}{\sqrt{(N-1)!}}\det|\varphi_a(1)\varphi_b(2)...\varphi_q(N-1)|. \tag{42}$$

The orbital $\varphi_i(L)$ is absent from the determinant on the right-hand side, and P is the sign of the permutation which carries the orbital $\varphi_i(L)$ over to the first position in the determinant on the left-hand side. If the determinant which a_i acts upon does not contain the orbital $\varphi_i(L)$ then the result of such an operation is zero. The creation and annihilation operators

for fermions satisfy the following anti-commutation formulas:

$$[a_i^\dagger, a_j^\dagger]_+ = 0, \tag{43}$$

$$[a_i, a_j]_+ = 0, \tag{44}$$

$$[a_i^\dagger, a_j]_+ = \delta_{ij}. \tag{45}$$

The product of an arbitrary number of creation and annihilation operators is called the second-quantized string. If the number of the creation operators in the string is the same as the number of the annihilation operators in it, the string is called particle conserving. It is obvious that any determinant may be put into correspondence with the string of creation operators, which acting on the vacuum vector creates this determinant.

Now we wish to establish the form of the operators in the Fock space which are in one-to-one correspondence with the operators in the Hilbert space. A general form of a particle-conserving operator C in the Fock space is

$$C = c_0 + \sum_{ij} c_{i;j}\ a_i^\dagger a_j + \sum_{ijkl} c_{ij;kl}\ a_i^\dagger a_j^\dagger a_l a_k + \sum_{ijklmn} c_{ijk;lmn}\ a_i^\dagger a_j^\dagger a_k^\dagger a_n a_m a_l + ... \tag{46}$$

It may be shown by simple manipulations that zero-particle operator, in other words a constant in the first quantization corresponds to the same constant in the second quantization. The one particle operator $\hat{C}(1)$ in the first quantization corresponds to the operator

$$C = \sum_{ij} c_{i;j}\ a_i^\dagger a_j \tag{47}$$

in the second quantization. Here $c_{i;j}$ are the matrix elements of $\hat{C}(1)$ in the first quantization:

$$c_{ij} = \langle \varphi_i | \hat{C}(1) | \varphi_j \rangle. \tag{48}$$

Analogously the two-particle operator $\hat{C}(1,2)$ in the first quantization will be written as

$$C = \sum_{ijkl} c_{ij;kl}\ a_i^\dagger a_j^\dagger a_l a_k \tag{49}$$

in the second quantization with the matrix elements

$$c_{ij;kl} = \langle \varphi_i(1)\varphi_j(2)|\hat{C}(1,2)|\varphi_k(1)\varphi_l(2)\rangle. \tag{50}$$

A lot of convenience of work in the second quantization is owing to the graphical representation of the operations with the creation and annihilation operators, which is generally known as diagrammatic techniques. There are numerous flavors of diagrammatic technique in quantum chemistry, and the one we are going to use in one of the chapters below allows to represent the operators in second quantization only if they are written in the so-called *normal form* or the normal order. The normal form of the second quantized string is denoted by the curly brackets $\{...\}$ and defined as the same string in which the minimal permutation of the operators has brought the creation operators to the left of the annihilation operators. The sign of the normally-ordered string is the sign of this minimal permutation:

$$\{a_i^\dagger a_j a_k ... a_l^\dagger a_m\} = (-1)^P a_i^\dagger a_l^\dagger ... a_j a_k a_m. \tag{51}$$

The *contraction* $[q]$ of two operators of second quantization (such as $a_i^\dagger$ and a_j) is the difference between their product and the normal form of their product:

$$A^{[q]}B^{[q]} = AB - \{AB\}. \tag{52}$$

Here q plays the role of a dummy index, showing which operator is contracted with which. For example, in $A^{[r]}B^{[q]}C^{[r]}D^{[q]}$ A is contracted with C, and B with D. Any fully contracted second-quantized string is a scalar since (52) is 0 in all cases except $a_i^{[q]}a_i^{\dagger[q]}$, when it is 1. One important result, commonly known as the *Wick's theorem*, states that the normal form of any second-quantized string may be written as the string itself minus the normal

forms of all possible contractions (single, double and so on) inside this string:

$$\begin{aligned}\{ABCD...EF\} = ABCD...EF &- \{A^{[q]}B^{[q]}CD...EF\} - \{A^{[q]}BC^{[q]}D...EF\} - ... \\ &- \{A^{[q]}B^{[q]}CD...EF\} - \{A^{[q]}BC^{[q]}D...EF\} - ... \\ &- \{A^{[q]}B^{[q]}C^{[r]}D^{[r]}...EF\} - \{A^{[q]}B^{[r]}C^{[r]}D^{[q]}...EF\} - ... \\ &- ... \\ &- \{A^{[q]}B^{[q]}C^{[r]}D^{[r]}...E^{[s]}F^{[s]}\} - \{A^{[q]}B^{[r]}C^{[s]}D^{[q]}...E^{[r]}F^{[s]}\} - ... \quad (53)\end{aligned}$$

A further insight into the structure and the properties of second-quantized operators may be obtained if we separate the space of our orbitals into 'occupied' and 'virtual' by introducing the Fermi level. For the N-electron system the N orbitals with the lowest orbital energies are called 'occupied'. The rest of the orbitals with higher energies (their number may be much larger than N) are called 'virtual'. Oftentimes the occupied orbitals are called 'holes' (and denoted by letter h) and the virtual ones are called 'particles' (denoted by letter p). The Fermi level is the energy of the highest occupied orbital.

The ground state is then written as

$$a_h^\dagger a_{h'}^\dagger ... a_{h''}^\dagger |0\rangle \equiv |\Phi_0\rangle, \quad (54)$$

where the indices h characterize the holes. Consequently, a singly-excited state, for example, is conveniently expressed as $a_p^\dagger a_h |\Phi_0\rangle$.

The operators a_h and $a_p^\dagger$ are the creation operators whereas $a_h^\dagger$ and a_p are annihilation operators. When a_h acts on $|\Phi_0\rangle$ it creates a hole (hence the name) among the occupied orbitals, and when $a_p^\dagger$ acts on $|\Phi_0\rangle$ it creates a particle on a virtual orbital. The annihilation operators $a_h^\dagger$ and a_p have a similar interpretation. Since we have chosen to separate our operators of second quantization into those having indices h and p, we must introduce a new definition for the normal ordering. We call this process 'renormalization'.

The normal ordering of the operators of second quantization with respect to the Fermi level is defined as follows. It is the string obtained by the minimal transposition executed

so that the operators of second quantization have the ordering a_h, $a_p^\dagger$, a_p, $a_h^\dagger$. The sign of the normally ordered string is defined by the sign of the transposition. For example,

$$\{a_p^\dagger a_h^\dagger a_{p'} ... a_{p''}^\dagger a_{h'}\} = (-1)^P a_{h'} a_p^\dagger a_{p''}^\dagger ... a_p a_h^\dagger. \tag{55}$$

2.6 The Goldstone diagrammatic technique

The Goldstone diagrammatic technique [38] which we are going to introduce in the present chapter rests on the method of second quantization described above. The diagrams depict the second quantized strings and operations on them and often greatly facilitate the work of a quantum chemist. Apart from the notational convenience that they bring, they are often essential for better understanding the structure of many complicated algebraic constructions.

If one desires to write a Goldstone diagram for a second-quantized operator, the latter must be preliminarily cast into the normal form. This task is easily accomplished by the application of the Wick's theorem (53). Figure 2.6 gives the details and various examples of the Goldstone diagrammatic notation.

There are three rows: each row contains several diagram with the corresponding algebraic formulas underneath. The black circle which one can see in some diagrams is called the *vertex*. It is used to distinguish the creation operators (or lines in the graphical notation) from the annihilation operators. The creation lines are always located above the vertex whereas, while the annihilation operators are found below it. Further, the particle lines face up whereas the hole lines face down. The first row in Figure 2.6 shows the four kind of lines which one can encounter in the diagrams. The index of the operator of second quantization is written next to the line. The second row shows various terms from the sums (47) and (49), when they are in the normal form. Note that the dashed line with the cross in the first two diagrams represent the one-particle operator while the zigzag line in the third diagram represents the two-particle operator. Apparently, the number of vertices in each diagram is equal to the number of particles the corresponding operator in the first quantization is

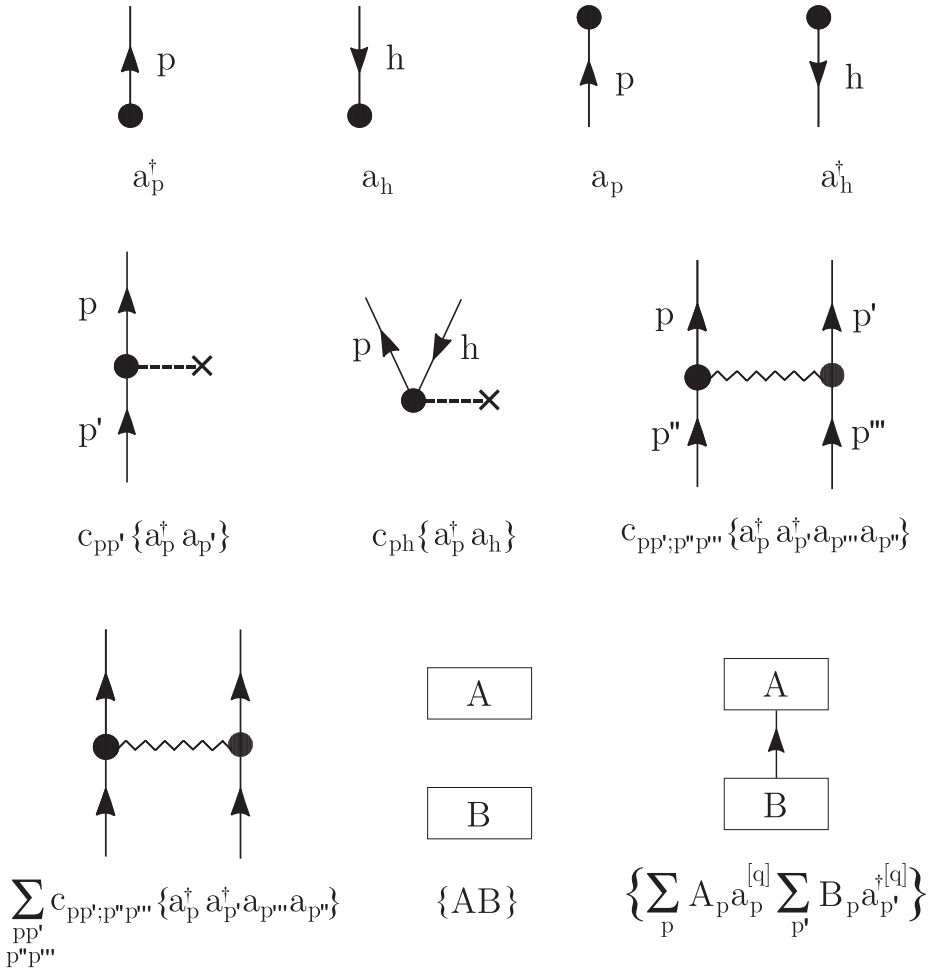


Figure 1: The Goldstone diagrammatic notation

associated with. The first diagram in the third row demonstrates that if there is a summation over some index then this index is not written in graphic notation (it is considered 'dummy'). The normal form of the product of two operators (without any contractions) is depicted as the diagram corresponding to the left operator over that corresponding to the right one. If there is a contraction between two operators (the third diagram in the third row) the corresponding lines are connected. The diagrammatic notation is constructed in such a way that the non-zero contractions $a_p^{[q]} a_p^{\dagger[q]}$ and $a_h^{\dagger[q]} a_h^{[q]}$ are always possible by joining the lines located directly one above the other. In all the cases of zero contractions the lines will face different directions or they will not be located directly one above the other

(they will be separated by the body of the operator).

CHAPTER III

HYBRID CORRELATION MODELS BASED ON ACTIVE-SPACE PARTITIONING: CORRECTING MP2 FOR BOND-BREAKING REACTIONS

3.1 Introduction

The restricted Hartree-Fock (RHF) theory provides a qualitatively incorrect description of homolytic bond-breaking reactions because it fails to include the additional electron configuration(s) which become degenerate when the fragments are far apart [4, 39]. Near-degeneracies of electron configurations can also complicate theoretical predictions for di-radicals, first-row transition metals, and other species. Unfortunately, popular post-Hartree-Fock theories which approximately account for electron correlation via singly and doubly-substituted Slater determinants are often incapable of fully overcoming the initial deficiencies of the Hartree-Fock wave function [40, 41, 42]. Although unrestricted Hartree-Fock (UHF) gives much more accurate energies than RHF at large inter-fragment separations, this comes at the price of massive spin contamination and the wave function no longer having the proper spin character. Moreover, correlated computations using unrestricted orbitals can yield less accurate energies in the intermediate bond-breaking region [41, 43, 44].

Although multi-reference methods are designed to handle near-degeneracies among electron configurations, they can be computationally costly and difficult to use for non-experts. Thus it is useful to assess the accuracy of single-reference methods for bond-breaking reactions and other examples of electronic degeneracies to see if some of them may be sufficiently accurate for certain applications. This can be done [45, 46, 47, 41, 43]

by comparing to full configuration interaction (full CI) results, which are numerically exact solutions to the electronic Schrödinger equation within the space spanned by the given one-particle basis set. At the same time, it is important to investigate possible improvements to single-reference methods which might make them more reliable for bond-breaking problems while retaining their simplicity and 'black box' character. Some efforts along these lines include the spin-flip approach of Krylov [48, 49, 50, 51] and the completely-renormalized coupled-cluster methods of Piecuch [52, 53], among others [54, 55, 56, 57, 58].

The simplest single-reference treatment of electron correlation is second-order Møller-Plesset perturbation theory (MP2). For many equilibrium properties, MP2 provides very reliable predictions which are suitable for all but the most demanding applications. Moreover, MP2 is the least expensive computationally of the conventional single-reference correlated methods. It is unfortunate, then, that RHF-based MP2 fails catastrophically for bond-breaking reactions or other cases where electron configurations become nearly degenerate. MP2 potential energy curves "turn over" at large inter-fragment distances, and the energy appears to diverge toward negative infinity. This behavior can be understood by examination of the spin-orbital energy equation for MP2 (assuming Hartree-Fock orbitals),

$$E^{\mathrm{MP2}} = E^{\mathrm{HF}} + \frac{1}{4}\sum_{ijab}\frac{|\langle ij||ab\rangle|^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}, \tag{56}$$

where E^{HF} is the Hartree-Fock reference energy, ϵ_p is the energy of orbital p, and $\langle ij||ab\rangle$ are the usual antisymmetrized two-electron integrals. As the bonding and antibonding orbitals of the bond being broken approach degeneracy at large inter-fragment separations, then at least one term in the above sum becomes an unphysically large negative number; namely, the term where i and j are the indices corresponding to the α and β spin components of the bonding orbital, and a and b are those of the two spin components of the antibonding orbital which becomes nearly degenerate with it. One might optimistically hope that MP2 could be 'fixed' for bond-breaking reactions by the replacement of this one

offending term. As we will see, such a simple remedy does not appear to be satisfactory. Nevertheless, a more careful examination of the failure of MP2 at large distances (analyzed via a decomposition of the energetic contributions using active-space concepts) suggests that one should replace the offending terms from MP2 with analogous terms from more robust theories; a minimal number of terms will be replaced so as to retain the fifth-order scaling of the conventional MP2 method.

One possibility is to replace the problematic MP2 terms with their counterparts from the more complete coupled-cluster theory. This is reminiscent of the strategy of using coupled-cluster theory for the more important 'strong pairs' and MP2 for the less important 'weak pairs' in local coupled-cluster theory [59, 60, 61]. Although coupled-cluster theory including single and double substitutions [62] (CCSD, the cost of which scales as the sixth power of the system size) is not itself foolproof for bond-breaking reactions, it is at least qualitatively correct in most cases where a single bond is broken — a vast improvement over MP2. Moreover, the errors in CCSD remain modest for some simple reactions [41]. Additionally, it seems likely that in some systems where the degeneracies are not too strong (e.g., certain transition metals), CCSD may yield reasonable results where MP2 would fail. Finally, the MP2 method can be regarded as a special case of CCSD under certain simplifying assumptions, making it theoretically appealing to consider replacing certain terms from MP2 with their CCSD counterparts. However, it is also clear that this would be one of the simplest approaches in a whole family of hybrid methods which could be made more robust via the inclusion of higher-order terms (via coupled-cluster or configuration interaction approaches).

3.2 Theory and discussion

The CCSD energy expression, again assuming Hartree-Fock orbitals, may be written as

$$E^{\mathrm{CCSD}} = E^{HF} + \frac{1}{4}\sum_{ijab}\langle ij||ab\rangle \left(t_{ij}^{ab} + 2t_i^a t_j^b\right) \tag{57}$$

It is clear that the MP2 energy expression can be obtained from the CCSD energy expression when single excitations are neglected and when the T_2 amplitudes are fixed at their first-order form $t_{ij}^{ab(1)} = \langle ij||ab\rangle/(\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b)$. This connection between MP2 and CCSD theory will be exploited in the present work.

Although it is well appreciated that the MP2 and CCSD methods can give rather different numerical results, especially for non-equilibrium geometries, it is conceivable that some terms constituting E^{MP2} and E^{CCSD} might be very similar. For instance, one might expect that the MP2 and CCSD theories describe the excitations from low-energy occupied to high-energy virtual orbitals equally well, whereas the excitations from highest occupied molecular orbital (HOMO) to lowest unoccupied molecular orbital (LUMO) may be handled with a significant difference in quality. Indeed, bond dissociation is accompanied by the HOMO and LUMO energies' gradually becoming degenerate, so that the denominator of the corresponding E^{MP2} term tends to zero, and at some point on the potential energy surface the perturbation expansion is no longer valid. However, the analogous term of E^{CCSD} does not exhibit divergent behavior even at larger bond distances. As the number of terms in the sums (56) and (57) is on the order of o^2v^2, where o is the number of doubly occupied orbitals and v is the number of virtual orbitals, it is impossible to compare the behavior of all the individual corresponding terms as a function of the bond length for a non-trivial molecule with a reasonable basis set. Instead, by using an appropriate active space, one can separate the occupied and virtual orbitals into restricted (R) and active (A) subsets, so that the correlation energy contributions to both (56) and (57) can be re-written as

$$\begin{aligned} E^Q_{corr} &= E^Q_{AAAA} + E^Q_{AAAR} + E^Q_{AARR} + E^Q_{ARAA} \\ &+ E^Q_{ARAR} + E^Q_{ARRR} + E^Q_{RRAA} + E^Q_{RRAR} + E^Q_{RRRR}, \qquad Q = \mathrm{MP2},\ \mathrm{CCSD}, \end{aligned} \tag{58}$$

so that only nine terms need to be compared. In the expression E^Q_{WXYZ} the first two lower indices (W and X) stand for the orbital subsets *from* which the electrons are excited, and

the last two lower indices (Y and Z) stand for the orbital subsets *to* which the electrons are excited. At this stage, the distinction between the restricted and active orbitals is purely notational: no actual restrictions were imposed on the excitations either from or to the R-orbitals.

In this initial study, we look at some simple bond-breaking reactions in BH, CH_4, and HF, where our previous full configuration interaction benchmarking studies show that CCSD works reasonably well [41]. For CH_4, the reaction considered is $CH_4 \rightarrow CH_3 + H$, where we have fixed the C–H distances and HCH angles at 1.086 Å and the tetrahedral value, respectively, for convenience [41]. In this study, we use the 6-31G* basis and correlate all electrons. For each molecule, we consider a minimal active space consisting only of the highest occupied and lowest unoccupied orbitals of the totally-symmetric irreducible representation, which at large distances will correspond to the bonding and antibonding orbitals for the bond being broken. For HF and CH_4, we also consider larger active spaces, chosen somewhat arbitrarily as ($6a_1$ $2b_1$ $2b_2$) and ($7a'$ $3a''$), respectively, where we give the number of active orbitals in each of the irreducible representations of the computational subgroup. All occupied orbitals except the core orbitals are made part of the active space. The results reported in this study were obtained using a simple spin-orbital coupled-cluster code which was easy to modify for our present purposes. Transformed integrals were obtained using the PSI3 package [63].

Figure (2) compares each E^{MP2}_{WXYZ} component to the corresponding E^{CCSD}_{WXYZ} component as a function of bond lengths for each of our test cases. The minimal active space results are of the greatest interest, since one may wonder whether only the $AAAA$ contributions to the total energy differ markedly. If the rest of the terms are described quantitatively (or at least qualitatively) in the same manner by both MP2 and CCSD theories, one could substitute the faulty E^{MP2}_{AAAA} term by the acceptable E^{CCSD}_{AAAA} term and in this way correct the total MP2 energy curve.

Let us discuss the minimal active space curves first. From Figure (2) we infer that,

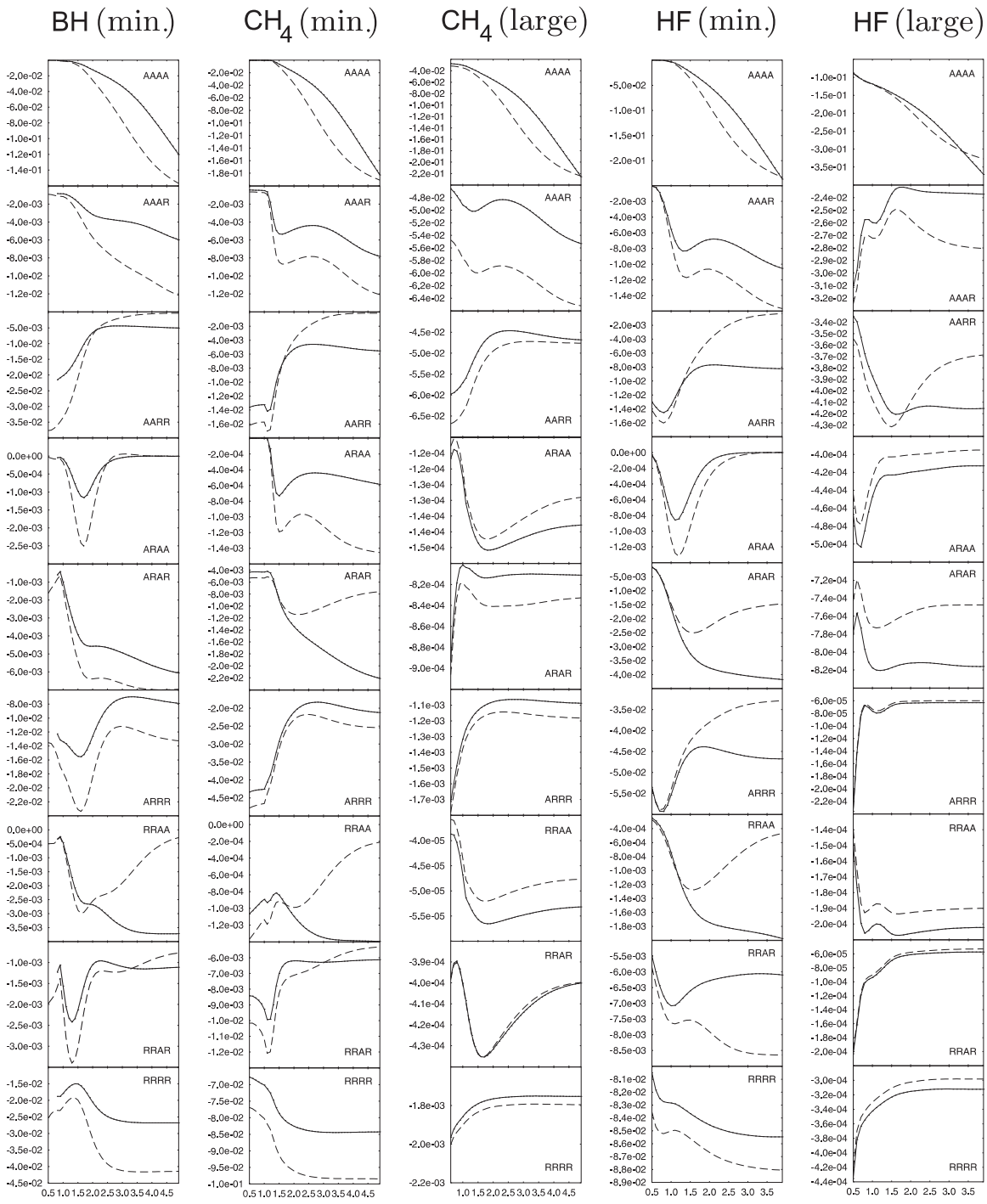


Figure 2: The orbital excitation components of MP2 and CCSD energies. The solid lines represent the MP2 components and the dashed ones correspond to the CCSD components. The WXYZ notation and the minimal and large active spaces are described in the text. Distances (x-axis) are in Ångstrom and energies (y-axis) are in atomic units.

indeed, the greatest quantitative difference between the $E^{\rm MP2}_{WXYZ}$ and $E^{\rm CCSD}_{WXYZ}$ contributions is in the case of the AAAA term. Both $E^{\rm MP2}_{AAAA}$ and $E^{\rm CCSD}_{AAAA}$ terms decrease with bond distance, but while the CCSD term levels out, the MP2 term goes down sharply, making the total MP2 energy apparently divergent. Unfortunately, the E_{AAAA} contribution is not the only one where discrepancies are observed between the MP2 and CCSD methods. Each pair ($E^{\rm MP2}_{WXYZ}$, $E^{\rm CCSD}_{WXYZ}$) has a significant quantitative difference (although much smaller than that for the $AAAA$ term), and some pairs exhibit even qualitative differences. For example, the behavior of $E^{\rm MP2}_{RRAA}$ and $E^{\rm CCSD}_{RRAA}$ curves in the case of BH molecule is strikingly dissimilar after about 2Å. The same observation concerns the $ARAR$ and $RRAA$ pairs of CH_4 as well as the $AARR$, $ARAR$, $ARRR$, $RRAA$, and $RRAR$ pairs of HF. It is pleasant to note though that for all the molecules studied the $RRRR$ pairs behave qualitatively in a similar way.

This analysis suggests that if we were to substitute only the $E^{\rm MP2}_{AAAA}$ term by the corresponding CCSD term in the MP2 energy expression,

$$E^{\rm MP2}_{AAAA} \longleftarrow E^{\rm CCSD}_{AAAA} \tag{59}$$

the resulting *modified* MP2 energy curve should avoid the disastrous turnover occasioned by the divergent $E^{\rm MP2}_{AAAA}$ term, but it might still exhibit some quantitiative or even qualitative errors. This conclusion is tested in Figure 2, which compares the total modified MP2 energy as obtained by eq. (59) and the conventional MP2 and CCSD energies for our three test cases, BH, CH_4, and HF.

Here, the term $E^{\rm CCSD}_{AAAA}$ has been computed simply by a conventional CCSD computation. We will call this energy the MP2+CCSD(CCSD) energy, where "MP2+CCSD" indicates that we have simply added the $E^{\rm CCSD}_{AAAA}$ energy component to the complementary MP2 energy components, and "(CCSD)" denotes how the term $E^{\rm CCSD}_{AAAA}$ was obtained (through a conventional CCSD computation). We see that the MP2+CCSD(CCSD) potential curves are vastly improved over the MP2 curves, and the results for BH and CH_4 seem to provide an excellent match to the full CCSD results [the curves appear to be parallel,

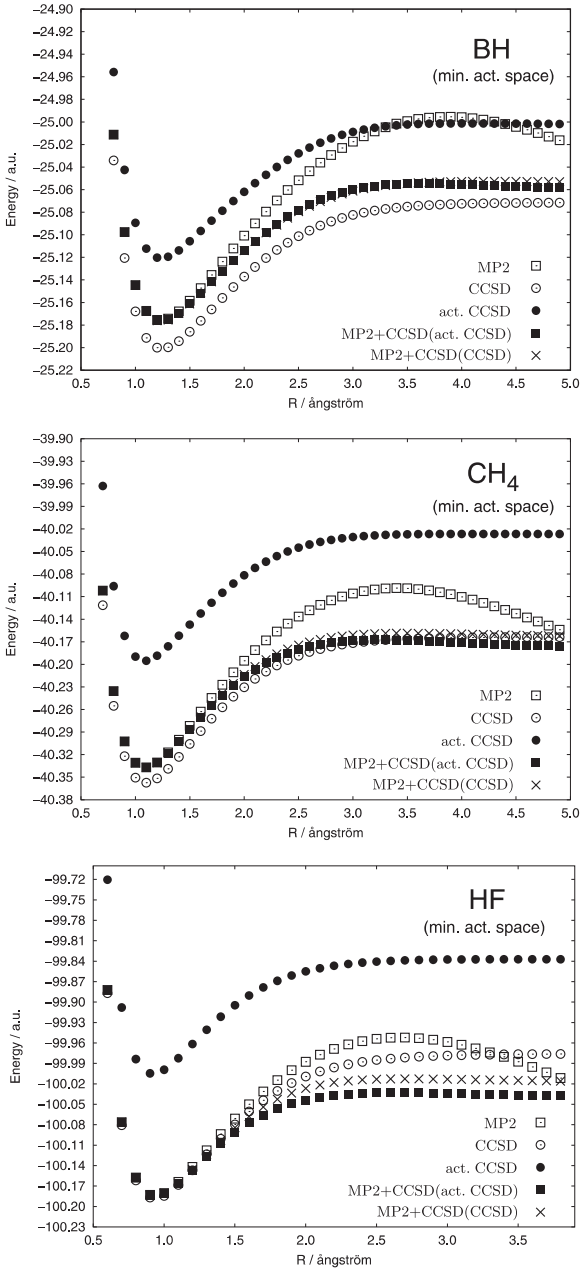


Figure 3: The performance of different hybrid MP2+CCSD methods for minimal active spaces.

with the MP2+CCSD(CCSD) curves shifted slightly higher in energy]. Closer scrutiny of the energies reveals a very slow decrease in the MP2+CCSD(CCSD) energies at large distances, and this downward drift is much larger and clearly visible on the graph for HF. The MP2+CCSD(CCSD) scheme, then, provides a great improvement over MP2 for these

cases, but it is not completely robust.

Now let us briefly discuss the results in case of the extended active space. In the extreme case when *all* orbitals are considered active, the modified MP2 energy becomes equal to the CCSD energy. Therefore it is reasonable to suppose that the gradual increase of the active space will provide a gradual transition from the MP2 energy to the CCSD energy. It is easy to see from the large active space curves presented in Figure (2) that now the components E^{MP2}_{WXYZ} only slightly differ from the components E^{CCSD}_{WXYZ} for almost all the components $WXYZ$ except the $AAAA$ term. The MP2+CCSD(CCSD) potential curves calculated corresponding to large active spaces and presented in Figure (4) also show a dramatic improvement in quality, and now they appear to be an excellent match for the CCSD curves for all three test cases.

The computational scheme discussed so far, MP2+CCSD(CCSD), is no less expensive than the CCSD method itself, because for the E^{CCSD}_{AAAA} value is obtained only after the CCSD iterations have converged. Therefore, in order to make this combined scheme practical we would need to approximate the E^{CCSD}_{AAAA} value by the energy obtained from a CCSD computation of a moderate cost, for example from a CCSD performed using active orbitals only. For a minimal active space, the cost of this CCSD computation would be negligible, and it would remain small so long as the number of active orbitals is small compared to the total number of orbitals. However, in using such an approximation to the E^{CCSD}_{AAAA} term, we introduce an additional error into the MP2+CCSD energy. We will call the energy of this less expensive procedure the MP2+CCSD(active CCSD) procedure, indicating that the term E^{CCSD}_{AAAA} is obtained from an active-space CCSD computation. Figures 3 and 4 demonstrate that this approximation to E^{CCSD}_{AAAA} makes the MP2+CCSD(active CCSD) potential curve significantly worse than the MP2+CCSD(CCSD) curve. Of course, the larger the active space the smaller the error of the E^{CCSD}_{AAAA} approximation, but the computational cost of making this approximation becomes larger. An additional problem is that the scheme ceases to be a 'black-box' method if active spaces other than the minimal active space are

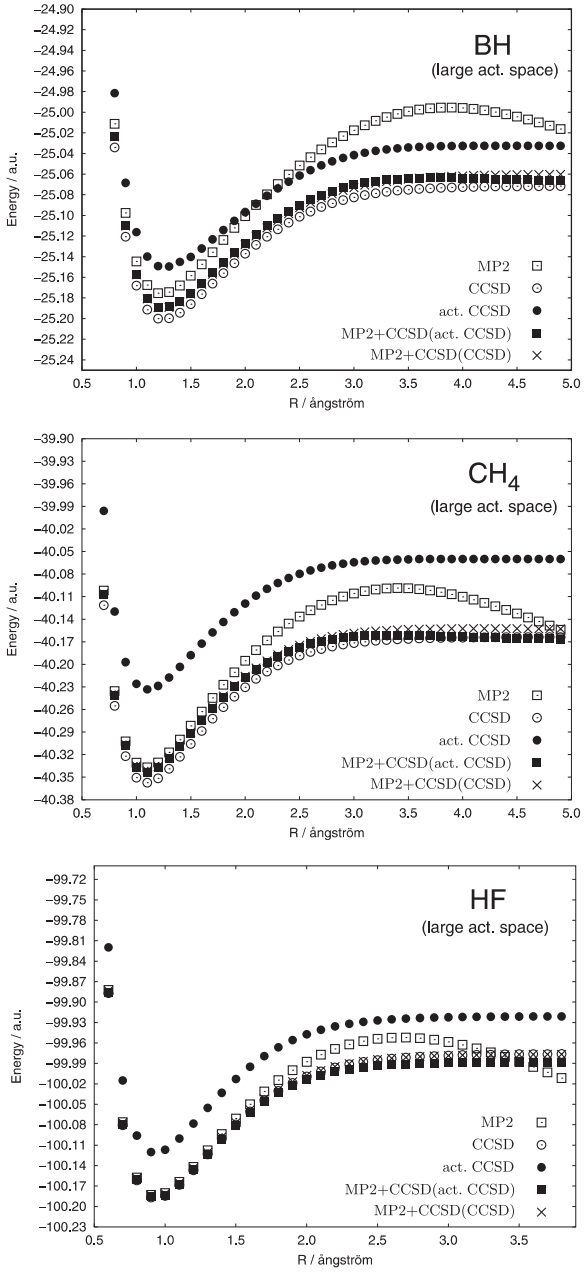


Figure 4: The performance of different hybrid MP2+CCSD methods for minimal active spaces.

used.

The results for MP2+CCSD(active CCSD) indicate that one cannot safely neglect the coupling between the active-only electron configurations and the other configurations. At

the same time, however, to obtain a practical advantage over conventional CCSD, the computation of the full CCSD wave function to obtain E_{AAAA}^{CCSD} must be avoided. How can these two requirements be satisfied? The most obvious solution is to solve for the coupled-cluster amplitudes giving rise to E_{AAAA}^{CCSD} *in the presence of the other amplitudes fixed in their MP2 form*. Indeed, one might hope that the explicit coupling of the coupled-cluster amplitudes to the MP2 amplitudes might improve results over the MP2+CCSD(CCSD) approach discussed above. We will denote this new alternative as MP2-CCSD, where the dash will indicate a coupling between the methods rather than a simple addition of terms obtained from separate computations. This form of hybrid MP2 and CCSD was previously investigated by Nooijen [64] in a different context; namely, as a promising way to reduce the cost of CCSD for computations of equilibrium properties and excited state energies. The general idea of reducing costs by restricting higher-order correlation terms to have at least a certain number of indices corresponding to active orbitals has ample precedent in coupled-cluster and configuration interaction methods (see, for example, Refs. [11, 65, 66, 67, 68, 55, 69, 70]).

It is easy to implement the coupled, hybrid MP2-CCSD approach. CCSD programs generally use the MP2 amplitudes as an initial guess in the iterative solution of the cluster amplitudes, and the MP2 energy is obtained in the first iteration of a CCSD procedure. A possible exception is the case of MP2 based on a restricted open-shell Hartree-Fock (ROHF) reference. First, there are several possible definitions of such theories, and second, the ROHF Fock operator is not diagonal unless a transformation to semicanonical orbitals is performed. See, for example, Ref. [71]. To implement the hybrid MP2-CCSD approach, one needs only to fix all amplitudes in their initial MP2 form except for those "internal" amplitudes t_{ij}^{ab} and t_i^a, all of whose indices belong to the active space orbitals. The final, converged energy may be written as

$$E_{corr}^{\mathrm{MP2-CCSD}} = E_{AAAA}^{\mathrm{MP2-CCSD}} + E_{AAAR}^{\mathrm{MP2}} + E_{AARR}^{\mathrm{MP2}} + E_{ARAA}^{\mathrm{MP2}} +$$

$$E_{ARAR}^{\mathrm{MP2}} + E_{ARRR}^{\mathrm{MP2}} + E_{RRAA}^{\mathrm{MP2}} + E_{RRAR}^{\mathrm{MP2}} + E_{RRRR}^{\mathrm{MP2}}, \tag{60}$$

where

$$E_{AAAA}^{\mathrm{MP2-CCSD}} = \frac{1}{4}\sum_{ijab}^{act}\langle ij||ab\rangle \left(t_{ij}^{ab} + 2t_i^a t_j^b\right), \tag{61}$$

assuming Hartree-Fock orbitals. The summation in the formula for $E_{AAAA}^{\mathrm{MP2-CCSD}}$ is restricted to the active orbital space.

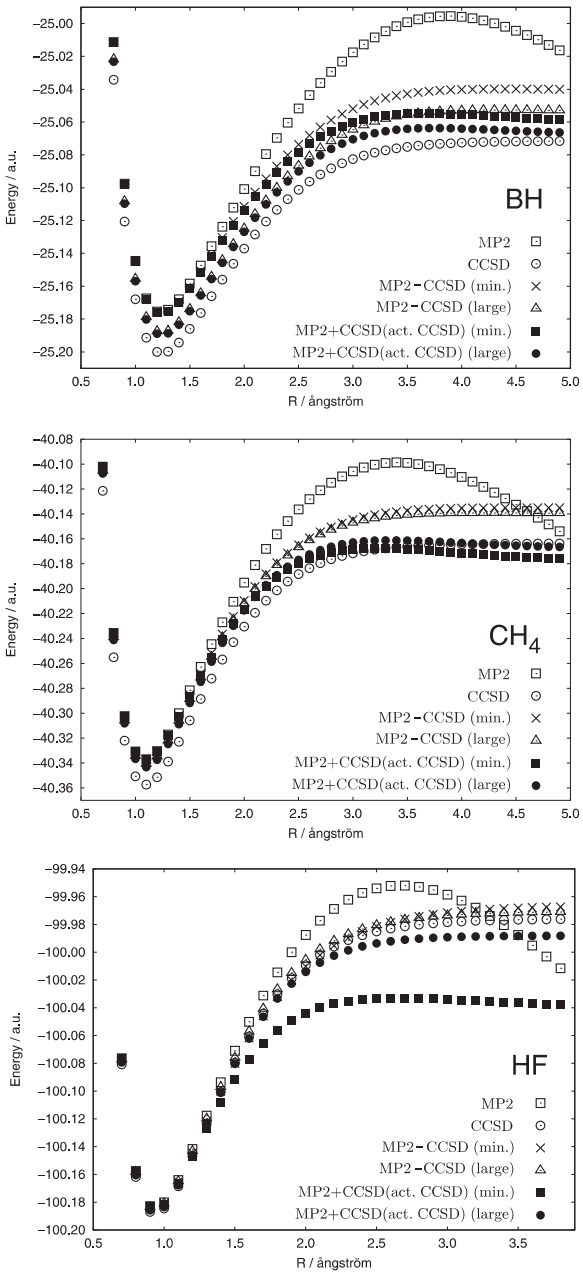


Figure 5: The performance of the coupled, hybrid MP2-CCSD method.

Figure 5 shows the performance of MP2-CCSD for various molecules and different active spaces. We may observe that even in the case of the minimal active spaces, the total energies derived from $E_{AAAA}^{\mathrm{MP2-CCSD}}$ are qualitatively correct for every test case; in contrast to some of the MP2+CCSD approaches discussed above, the tendency of the potential curves to bend downward at large distances has been eliminated by the coupling of the coupled-cluster amplitudes to the MP2 amplitudes. At bond lengths close to equilibrium, the MP2-CCSD curves follow the MP2 curves very closely, but at larger separations, when the MP2 energies start to behave incorrectly, the MP2-CCSD curves remain almost parallel to the CCSD total energy curves and at small and intermediate distances fall between the MP2 and CCSD curves. Increasing the size of the active space improves the picture even further.

It is easy to estimate the cost of such a coupled, hybrid method. In the notation of Stanton and co-workers, [72] the most time-consuming step of the CCSD procedure is the evaluation of the $\mathcal{W}_{abef} \rightarrow \Delta t_{ij}^{ab}$ contribution, which includes the four-virtual $\langle ab||ef\rangle$ components. This step normally scales as o^2v^4, where o and v are the number of occupied and virtual orbitals, respectively. In the MP2-CCSD approach, the cost reduces to $v^2O^2V^2$, where O and V represent the number of active occupied and active virtual orbitals, respectively. In the minimal active spaces considered here, $O = V = 1$, and typically O and V will be of order 1. The cost of forming the $\mathcal{W}_{mbej}$ intermediate is reduced from o^3v^3 to o^2v^2OV, and the cost of forming the $\mathcal{W}_{mnij}$ intermediate is reduced from o^4v^2 to $o^2v^2O^2$. For a minimal active space, then, the computational scaling is no worse than that of MP2 theory, and for larger active spaces, the cost is still considerably less than that of conventional CCSD.

In order to investigate the errors of these hybrid methods more carefully, we present errors versus full CI energies for the HF molecule in Figure 6.

These are not the same full CI results as in our previous benchmarking study Ref. [41]; that study used the larger 6-31G** basis and froze core orbitals. The full CI results may be

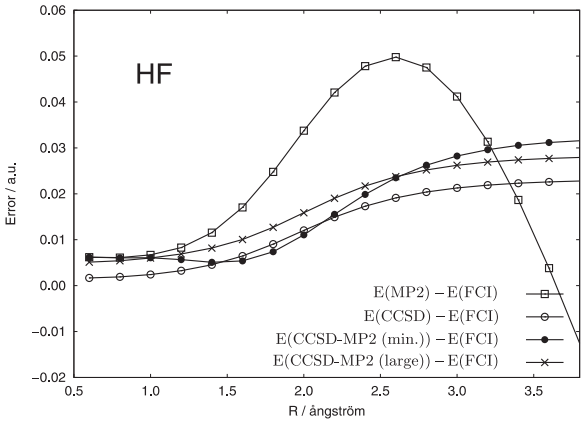


Figure 6: Errors for the coupled, hybrid MP2-CCSD method and for conventional MP2 and CCSD compared to full configuration interaction for the HF molecule.

considered numerically exact solutions of the electronic Schrödinger equation within the given basis set. One wishes for flat error curves, which indicate that an approximation is yielding a potential curve parallel to the full CI curve (a mere shift in the potential, giving a constant error across the surface, is of no chemical significance). The figure reveals that the MP2-CCSD error curve tracks that of CCSD extremely well when the large active space is used. For the minimal active space, the MP2-CCSD error curve follows the general shape of the CCSD error curve at short and long distances, but the errors happen to dip slightly below the CCSD errors at intermediate distances. Both the small and large active space MP2-CCSD error curves are tremendously improved over those of conventional MP2.

3.3 Conclusions

We have analyzed the failure of RHF-based MP2 for bond-breaking reactions in terms of energetic contributions from different orbital subspaces. The one double excitation from the sigma bonding orbital of the bond being broken to its antibonding counterpart accounts for the majority of the error in MP2 at large interfragment separations due to the energy denominator becoming nearly zero. Using more robust methods such as coupled-cluster

theory to replace this one term vastly improves results, but there remains a slight, unphysical drift downward in the energy at large separations due to some other energy components involving the active orbitals. This residual problematic behavior can be eliminated by coupling the solution of the active-space coupled-cluster problem to the MP2 amplitudes for the remaining excitations. This results in a method with a computational cost similar to that of MP2 (for minimal active spaces) but which behaves qualitatively correctly for reactions in which a single bond breaks, the error being roughly comparable to that observed for the much more complete CCSD method. Results of CCSD quality can be systematically approached using larger active spaces for the CCSD part of the CCSD-MP2 hybrid wave function.

Many other approaches to bond-breaking reactions and near-degeneracies of electron configurations begin with a robust active-space computation (e.g., complete-active-space self-consistent-field, CASSCF) and then proceed to add a description of electron correlation involving non-active (restricted) orbitals. Here, we have followed something of a reversed procedure, fixing the description of electron correlation involving restricted orbitals at the MP2 level and then proceeding to solve for a more robust description of the active space. It is perhaps surprising that this approach works as well as it does for the cases considered. It is encouraging that the all-restricted ($RRRR$) energy component appears to be estimated well by MP2 even in the presence of nearly degenerate electron configurations, but we did notice some differences between MP2 and CCSD for some of the other components which we nevertheless fixed in their MP2 form. The coupling of the active-space coupled-cluster amplitudes to the fixed MP2 amplitudes allows them to respond in such a way as to produce qualitatively correct potentials.

The current results suggest that even in the presence of nearly degenerate electron configurations, MP2 theory may remain useful for estimating correlation coefficients except for those in which all indices correspond to active orbitals. In future work, we will investigate this possibility more fully by examining additional bond-breaking situations. Clearly,

MP2-CCSD will not work for breaking double or triple bonds, because CCSD itself fails in these situations. However, one can easily imagine more sophisticated approaches which build upon the ones investigated here.

The material of this chapter was published in the *Journal of Chemical Physics* [73].

CHAPTER IV

HYBRID CORRELATION MODELS BASED ON ACTIVE-SPACE PARTITIONING: SEEKING ACCURATE AB INITIO METHODS FOR BOND-BREAKING

4.1 Introduction

The accurate description of potential energy surfaces (PES's) has been and still remains one of the primary objectives of quantum chemistry [42]. Unfortunately, the standard hierarchy of single-reference electron correlation methods does not work reliably for bond-breaking reactions, particularly for reactions which make or break multiple bonds. The standard flavors of density functional theory (DFT) are usually unsuitable for the computation of PES's due to their poor qualitative and quantitative performance, especially at stretched geometries [41, 74]. Although a variety of multireference methods can provide accurate results in virtually any bond-breaking reaction, in practice they tend to be difficult to derive, implement, and use, and moreover they can be very expensive computationally. Thus it remains desirable to investigate more "black box" bond-breaking methods with favorable computational scaling. Recent work along these lines includes new methods by Head-Gordon and co-workers based on ideas from the generalized valence bond perfect-pairing approach [75, 76, 77], the spin-flip approach of Krylov [48, 78, 49, 51, 50], and the method of moments and completely-renormalized coupled-cluster methods of Piecuch [52, 53, 79, 80]. In this work, we explore hybrids of coupled-cluster and perturbation theories for reactions breaking single bonds.

Among the standard quantum chemical methods based on the restricted Hartree-Fock (RHF) reference, the cheapest qualitatively correct method for breaking single bonds in the ground state is the coupled cluster theory with the inclusion of single and double excitations (CCSD) [62, 72]. Its formal scaling with the total number of occupied (o) and virtual (v) molecular orbitals and the number of iterations $N_{\rm it}$ required to converge the nonlinear CCSD equations is $N_{\rm it}o^2v^4$. When we refer to the quality of the method in relation to bond breaking we mean the correctness of the shape of the potential energy curve produced by this method rather than the absolute error in energy. The CCSD energy curves for reactions breaking single bonds usually overestimate the dissociation energy but they are smooth and devoid of artifacts such as divergence at large interatomic distances. The latter defect is only too common among the methods which utilize the perturbation theory: for example, both the second-order Møller-Plesset (MP2) theory and the CCSD(T) method [81] (often referred to as the "golden standard" of quantum chemistry) fail catastrophically at non-equilibrium geometries [41, 82]. The failure of MP2 is especially regrettable since this method has a very low computational scaling, $\mathcal{O}(N^5)$, where N is the total number of orbitals, $N = o + v$. Another method that has a low formal scaling, $N_{\rm it}N^5$, is the CC2 method of Christiansen *et al.* [83] Unfortunately, its behavior at large interatomic distances remains largely unexplored (see, however, studies of CC2 energy curves around equilibrium geometries in Refs. [84, 85]). We touch on this topic in the current study.

Alternatively, when standard single-reference methods are used in conjunction with unrestricted Hartree-Fock (UHF) orbitals, the divergence at large interatomic distances is remedied. However, the UHF-based correlated wave functions often suffer from the serious spin contamination [86, 87]. Besides, the potential energy curves obtained by these methods can display an erroneous behavior in the intermediate bond-breaking region. In the case of UMP2 this erroneous behavior sometimes becomes so pronounced that it may be regarded as a grave defect of the method [41].

Thus, it appears that one has to tolerate the computational scaling of $N_{\rm it}N^6$ or higher in

order to study the bond-breaking processes with at least qualitative correctness. Recently [73], we proposed a very simple computational scheme which scales as N^5 but approaches the CCSD method in accuracy. This method, which we originally denoted as MP2-CCSD [in this chapter we refer to it as MP2-CCSD(I)], is a hybrid between the MP2 and the CCSD theories and benefits from the scaling of the former and the accuracy of the latter. It relies on the orbital partitioning into active and restricted spaces, which might seem unfortunate in that the user must choose which orbitals to make active. However, we verified that even in the case of the minimal active spaces (which can often be determined *a priori*) our method performs in a very satisfactory manner and is a vast improvement over MP2. With the modest increase of the size of the active space (which does not deteriorate the favorable N^5 scaling) the potential energy curves generated by MP2-CCSD(I) become essentially parallel to those generated by CCSD. In this work we describe and test a new $\mathcal{O}(N^5)$ hybrid method MP2-CCSD(II) which is similar to MP2-CCSD(I) in structure but is significantly more accurate so that it rivals CCSD in accuracy even when the minimal active spaces are used. In section II we present the methodology behind the MP2-CCSD(I) method and in section III we give the description of the MP2-CCSD(II) method. Illustrative results are presented in section IV.

4.2 The hybrid methodology

For RHF or UHF orbitals, the correlation energy of MP2 theory is written as the sum over all possible double excitations as in (37). It is always possible to formally divide the orbital space into four disjoint subsets: occupied active, occupied restricted, virtual active and virtual restricted orbitals [see Figure 7(a)]. Note that, so far, the denominations 'restricted' (R) and 'active' (A) do not indicate any constraint on the orbital excitations – these names are simply used for the notational convenience.

Any double excitation from the closed-shell reference shown symbolically in Figure 7(a) may be then labelled by the four-letter code $WXYZ$ where the first two letters (W and X)

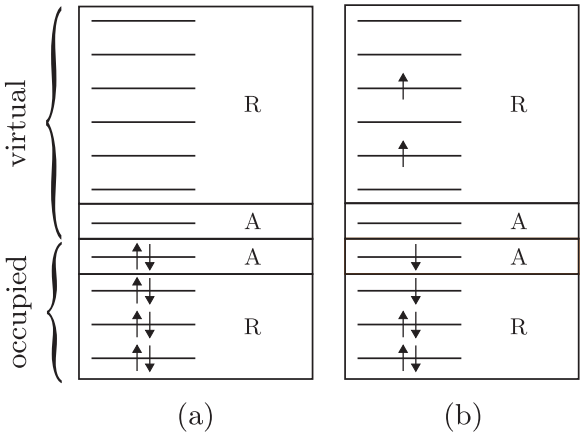


Figure 7: (a) The separation of the orbital space into four subspaces. (b) An example of our notation: ARRR-type excitation.

stand for the subspaces (A or R) from which the excitation was made and the last two letters (Y and Z) indicate the subspaces to which the electrons were excited. Obviously, $WXYZ$ is equivalent to $XWYZ$ etc. As an example, Figure 7(b) shows a ARRR-type excitation. In a similar manner, any single excitation may be labelled by the two-letter code WX where W shows from which subspace the electron was excited and X to which space it was excited.

Using this notation, we may rewrite the energy expression for the MP2 correlation energy as the sum of nine contributions:

$$E_{\rm corr}^{\rm MP2} = E_{AAAA}^{\rm MP2}+E_{AAAR}^{\rm MP2}+E_{AARR}^{\rm MP2}+E_{ARAA}^{\rm MP2}+E_{ARAR}^{\rm MP2}+E_{ARRR}^{\rm MP2}+E_{RRAA}^{\rm MP2}+E_{RRAR}^{\rm MP2}+E_{RRRR}^{\rm MP2}. \tag{62}$$

We note that the CCSD spin-orbital energy expression,

$$E_{\rm corr}^{\rm CCSD} = \frac{1}{4}\sum_{ijab}\langle ij||ab\rangle \left(t_{ij}^{ab} + 2t_i^a t_j^b\right), \tag{63}$$

reduces to the MP2 spin-orbital expression (37) in the event that single excitations are neglected ($t_i^a = 0$) and the doubles amplitudes are fixed in their first-order form, $t_{ij}^{ab(1)} = \langle ij||ab\rangle/(\epsilon_i+\epsilon_j-\epsilon_a-\epsilon_b)$. Indeed, the MP2 energy is given as the first iteration of the CCSD procedure for RHF or UHF orbitals when the MP2 guesses are used for the amplitudes.

This close connection between MP2 and CCSD is exploited in the current study.

In our previous work [73] we demonstrated for several small molecules that the $AAAA$ term, comprising no more than a handful of excitations for small active spaces, is primarily responsible for the divergence of the MP2 energy at large interatomic separations. The mechanistic substitution of the E^{MP2}_{AAAA} term with the E^{CCSD}_{AAAA} term (obtained from the CCSD calculation either in the full or active orbital space) in (62), which we called MP2+CCSD, does not lead to a very satisfactory potential energy curve, although even this simple operation redresses the sharp divergence of the MP2 energy curve. The MP2+CCSD energy curves show a small but nevertheless noticeable 'turning over' at large interatomic distances which is clearly not a physical behavior. The recipe for the proper replacement of E^{MP2}_{AAAA} is to do so *self-consistently*, that is adjust the $AAAA$-contribution in the presence of other contributions. This idea lies in the foundation of hybrid methods previously developed by Nooijen [64] for the investigation of excited states. The method that employs this methodology includes the following simple steps:

(i) Set up a CCSD calculation using the MP2 t^a_i and t^{ab}_{ij} amplitudes as a guess: $t^a_i = 0$ and $t^{ab}_{ij} = \langle ij||ab\rangle/(\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b)$.

(ii) Procede with solving the CCSD equations but update only the those single and double t-amplitudes that involve excitations within the active space only.

(iii) Terminate the iterations when the active space amplitudes and energy no longer change.

This approach which we called MP2-CCSD in our previous paper [73] will be referred to as MP2-CCSD(I) here. The convergence of this procedure is usually no worse than the convergence of the conventional CCSD equations. If the typical dimension of the active space is on the order of just a few orbitals (σ and σ^* for the minimal active space), then step (ii) has the computational expense around $\mathcal{O}(N^4)$. The next section introduces an improved MP2-CCSD(II) method and gives details as to the scaling of the intermediates which are computed on each iteration step. In summary, the cost of the MP2-CCSD(I) method is dominated by the atomic orbital (AO) to molecular orbital (MO) transformation

and is $\mathcal{O}(N^5)$. The potential energy curves generated by the MP2-CCSD(I) method normally level off at stretched geometries and show qualitative and quantitative advantages over the MP2+CCSD curves and dramatic improvements over simple MP2.

4.3 An improved $\mathcal{O}(N^5)$ hybrid method

The computational advantage gained in the MP2-CCSD(I) and MP2-CCSD(II) methods over the conventional CCSD method may be better understood from the analysis of the CCSD equations. We do not wish to complicate the present discussion and will therefore work with the CCSD equations written in a schematic form which stresses the most salient points. For greater detail, we refer the interested reader to the paper by by Stanton and co-workers [72] on the efficient implementation of CCSD.

The equation for each single-excitation t_1-amplitude may be written as a function f_1

$$t_i^a = f_1 \left(\sum_e t_i^e \sum_{mnf} t_{mn}^{af} \langle mn||ef\rangle, \sum_m t_m^a \sum_{nef} t_{in}^{ef} \langle mn||ef\rangle, \right.$$
$$\left. \sum_{mef} t_{im}^{ef} \langle ma||ef\rangle, \sum_{men} t_{mn}^{ae} \langle nm||ei\rangle, ... \right) \tag{64}$$

in which we explicitly mention as arguments only the most computationally expensive terms. The cost of the terms in parentheses are o^2v^3, o^3v^2, o^2v^3, and o^3v^2, respectively. Here o is the number of the occupied orbitals and v is the number of virtual orbitals. Similarly, the equation for each double-excitation t_2-amplitude is of the form:

$$t_{ij}^{ab} = f_2 \left(\sum_{mn} t_{mn}^{ab} \sum_{ef} t_{ij}^{ef} \langle mn||ef\rangle, \sum_{ef} t_{ij}^{ef} \langle ab||ef\rangle, \right.$$
$$\left. \sum_{me} t_{im}^{ae} \sum_{nf} t_{jn}^{fb} \langle ma||ef\rangle, ... \right) \tag{65}$$

The terms in the parentheses scale as o^4v^2, o^2v^4, o^3v^3, respectively. When we update only the AA t_1 amplitudes the arguments of the function f_1 reduce to the scalings ov^2OV, o^2v^2O, ov^2OV, and o^2vOV, respectively, where O is the number of occupied orbitals in

the active space and V is the number of virtual orbitals in the active space. Typically, $O \approx V \sim 1$, so we may think of these scalings as ov^2, o^2v^2, ov^2, and o^2v, respectively. In all the subsequent estimations we make the assumption that O and V do not grow with the size of the basis set. Updating only the $AAAA$ terms when we solve the MP2-CCSD equations reduces the scaling of the arguments of the function f_2 to $o^2v^2O^2$, $o^2v^2V^2$, o^2v^2OV, respectively. Or, removing the O and V dependencies, we obtain the identical scaling o^2v^2 for each term. Taking into account the iterative nature of the method, we conclude that the computational cost of the MP2-CCSD(I) procedure is $(N_{\mathrm{it}}o^2v^2)$. It is reasonable to assume that N_{it} does not depend on the size of the system, and so the total cost of the MP2-CCSD method is dominated by the orbital transformation procedure. Thus, the MP2-CCSD(I) has the formal scaling $\mathcal{O}(N^5)$.

It is easy to notice, however, that updating certain other types of amplitudes together with the AA and $AAAA$ amplitudes increases the cost of the resulting hybrid method only marginally and still keeps it much lower than the cost of the regular CCSD method. If we update the AR, RA, and RR single-excitation amplitudes, the cost of this operation will scale as o^2v^3. Further, if we also update the $ARAA$ and $AAAR$ double-excitation amplitudes, the worst scaling that will result from this operation will be o^2v^3V, or simply o^2v^3 if V is on the order of 1. The method in which we update the AR, RA, RR, $AAAA$, $ARAA$ and $AAAR$ amplitudes self-consistently in the presense of the rest of the amplitudes computed by the MP2 method we call the MP2-CCSD(II) method, scaling as $\mathcal{O}(N_{\mathrm{it}}o^2v^3)$. With respect to the increase of the size of the system (if N_{it} is assumed constant), the scaling of MP2-CCSD(II) is still not worse than that of the MP2 method.

In constructing the MP2-CCSD(I) and MP2-CCSD(II) methods we relied on the MP2 theory as a source of inexpensive t-amplitudes. One might ask whether there exists some other choice of the low-cost method. The Epstein-Nesbet (EN) pair-correlation theory or related constructs, whose computational cost is dominated by the AO-MO transformation,

is worthy of investigation in this respect. We utilized the following formula for the computation of the double excitation amplitudes:

$$t_{ij}^{ab} = \frac{\langle ij||ab\rangle}{e_{ij} - \langle \Psi_{ij}^{ab}|H - E_0|\Psi_{ij}^{ab}\rangle}, \tag{66}$$

where E_0 is the Hartree-Fock energy and e_{ij} are pair energies

$$e_{ij} = \sum_{a<b} \langle ij||ab\rangle t_{ij}^{ab} \tag{67}$$

which constitute the correlation energy:

$$E_{\text{corr}} = \sum_{i<j} e_{ij}. \tag{68}$$

Equations (66-67) are solved iteratively until the values t_{ij}^{ab} and e_{ij} no longer change. We call this approach TCEPA (truncated coupled electron pair approximation) because its formulas naturally arise from the truncation of a summation in the well-known CEPA equations [88, 89]:

$$\langle \Psi_{ij}^{ab}|H|\Psi_0\rangle + \sum_{\substack{k<l \\ c<d}} \langle \Psi_{ij}^{ab}|H - E_0|\Psi_{kl}^{cd}\rangle t_{kl}^{cd} = e_{ij} c_{ij}^{ab}, \tag{69}$$

and e_{ij} is

$$e_{ij} = \sum_{c<d} \langle \Psi_{ij}^{cd}|H|\Psi_0\rangle t_{ij}^{cd}. \tag{70}$$

Observe that the neglect of e_{ij} in the denominator of (66) brings us to the second order EN perturbation theory (which is equivalent to EN pair-correlation theory), and further approximation of $\langle \Psi_{ij}^{ab}|H - E_0|\Psi_{ij}^{ab}\rangle$ through $\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b$ yields the MP2 theory. Some denominators in the EN perturbation theory approach zero as the bond is being broken. This may be explained by the fact that certain orbitals i and a (as well as j and b) necessarily become degenerate along the dissociation coordinate and the expression $\langle \Psi_{ij}^{ab}|H|\Psi_{ij}^{ab}\rangle$ approaches E_0. A few computations convinced us that the EN perturbation theory diverges even faster than MP2. Murray and Davidson [90], who compared the MP theory with one of the flavors of the EN theory for equilibrium geometries and up to the fifth order in the perturbation, also arrived at the conclusion that MP gives more predictable energies. TCEPA, however,

promises a better dissociation behavior than the regular EN perturbation theory. If e_{ij} remains in the denominator (as in TCEPA) then the denominator is not likely to turn into zero since e_{ij} is the part of the correlation energy which actually becomes constant at the end of the dissociation. We also constructed the hybrid TCEPA-CCSD(I) and TCEPA-CCSD(II) models built exactly after the MP2-CCSD(I) and MP2-CCSD(II) models, respectively (the types of the amplitudes updated are the same). In TCEPA-CCSD(I) and TCEPA-CCSD(II) the t-amplitudes which are not updated in the course of solving the CCSD equations come from equations (66-67). Observe that by combining TCEPA with CCSD we do not attempt to correct or improve some particular feature of TCEPA (as we did it with MP2 by substituting its AAAA amplitudes with the CCSD amplitudes). We merely wish to describe as many amplitudes as possible by a higher-quality method (CCSD) without disturbing the computational scaling of the lower-quality method (TCEPA).

One more $\mathcal{O}(N^5)$ candidate for a possible hybridization with coupled cluster method is CC2. As demonstrated below, the divergence of CC2 at large interatomic separations is even worse than that of MP2, and therefore we ruled out the idea of constructing a hybrid method built upon CC2.

4.4 Results and discussion

The hybrid methods introduced in the previous chapters were implemented in a prototype code built upon the PSI3.2 [63] suite of quantum chemical programs and libraries. Here we test the performance of the theoretical constructs discussed above against results from full configuration interaction (FCI), which exactly solves the nonrelativistic electronic Schrödinger equation within the given one-particle basis set. Because of the high cost of these FCI computations, we restrict our attention in this work to the small molecules H_2, BeH^+, BH, HF, H_2O, CH^+, CH_4, and Li_2, all in the 6-31G* basis set. We note that the CCSD method, which the MP2-CCSD hybrids attempt to mimic at lower computational cost, performs reasonably well for bond-breaking in these molecules [41]. The active

spaces of the first five molecules consist of just σ and σ^*, whereas those of CH^+, CH_4, and Li_2 are slightly larger (because of the energetic proximity of other orbitals to σ and σ^* or the intersection of σ and σ^* with other orbitals along the dissociation curve). The active spaces of CH^+, CH_4, and Li_2 are $(2a_1\, b_1\, b_2)$, $(3a'\, a'')$, and $(2a_g\, b_{1u}\, b_{2u}\, b_{3u})$, respectively. We considered the reactions in which a single bond to hydrogen is broken, or in the case of Li_2, the unimolecular dissociation $Li_2 \rightarrow 2Li$. In the case of CH_4, for simplicity we fixed the non-dissociating bonds at 1.086 Å, and the HCH angles were 109.47122 degrees. Likewise in H_2O, one bond length was fixed at 0.967 Å, and the HOH angle was 107.6 degrees.

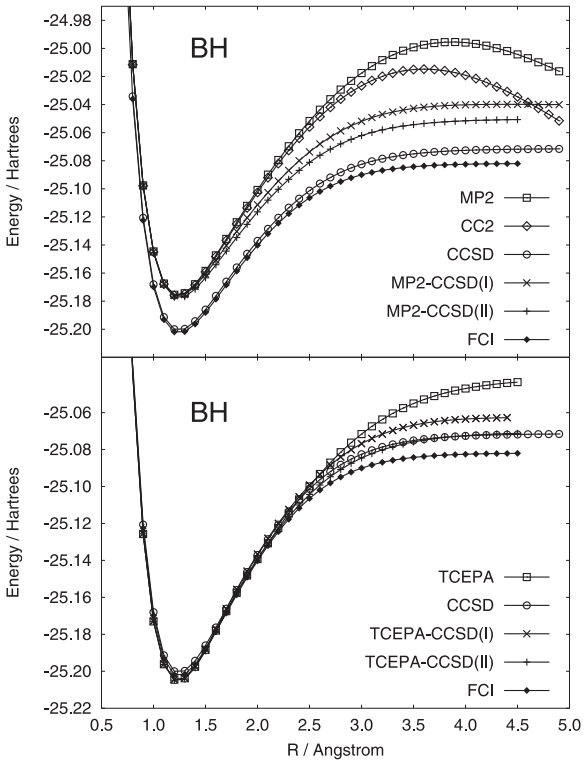


Figure 8: Performance of the hybrid theories on the BH molecule in 6-31G* basis set.

Figures 8-10 demonstrate the potential energy curves obtained with our hybrid models, which are plotted together with those obtained with the traditional models MP2 and CCSD, as well as with the CC2 method and with FCI. Potential energy curves for the other test cases are qualitatively similar. Let us initially consider the more standard methods, plus the

MP2-CCSD hybrid methods, shown in the top half of the figures. The standard methods MP2 and CCSD feature their typical behavior in Figures 8 (BH) and 9 (CH_4).

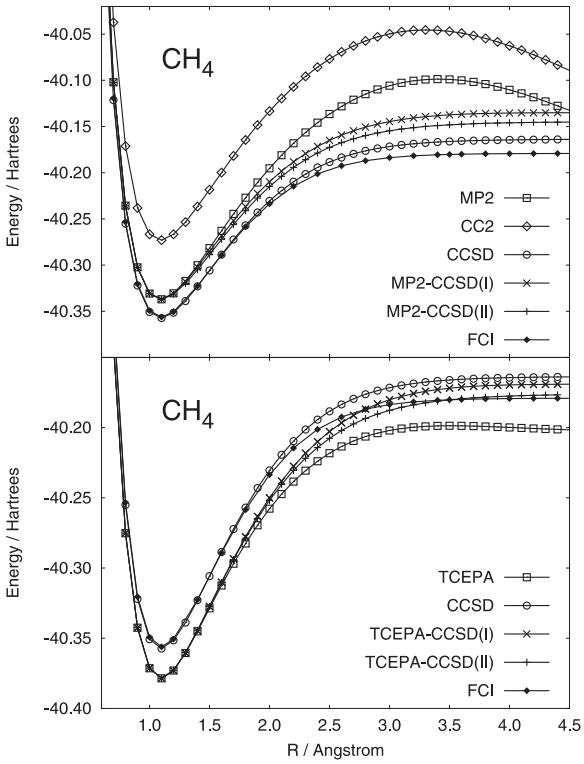


Figure 9: Performance of the hybrid theories on the CH_4 molecule in 6-31G* basis set.

The MP2 curve becomes unphysical at intermediate internuclear distances and then the energy diverges at large internuclear distances. Note that the CC2 curves are qualitatively similar to MP2 and also show divergence at large distances. CCSD, in contrast, performs reasonably well, yielding an energy which is somewhat too high at the dissociation limit. Our initial MP2-CCSD hybrid method, MP2-CCSD(I), behaves like MP2 near equilibrium but avoids the unphysical behavior of MP2 at larger internuclear distances. Moreover, the MP2-CCSD(I) curves are roughly parallel to those of CCSD, which the method approximates. The same can be said for the MP2-CCSD(II) method, which is quite similar to MP2-CCSD(I) but yields somewhat lower energies at large distances. A more quantitative comparison of these methods is presented below.

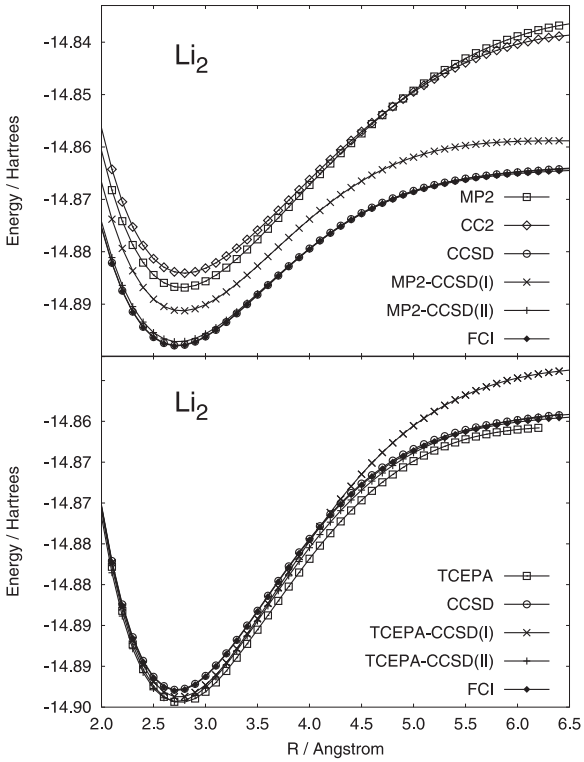


Figure 10: Performance of the hybrid theories on the Li_2 molecule in 6-31G* basis set.

The performance of the methods is somewhat different for the dissociation of Li_2, shown in Figure 10. The behavior of MP2 is again unphysical (and the curve turns over again at larger distances than those shown in the figure). Again, CC2 behaves similarly to MP2. However, in this case the CCSD curve is nearly coincident with FCI. MP2-CCSD(I) again produces a reasonable curve, although it yields a significantly lower energy than MP2 near equilibrium in this case. In contrast to its behavior for BH and CH_4, MP2-CCSD(II) now gives a much lower energy than MP2-CCSD(I) at all distances and indeed is nearly identical to CCSD and FCI.

Let us now turn to the lower parts of Figures 8-10, which display results from the TCEPA method introduced above. We had noted that the energy denominators in (66) are non-vanishing, and so one might expect better behavior at large internuclear separations

than that seen for standard MP2. Indeed, the behavior of TCEPA is perhaps not as catastrophic as MP2, but the curves are seriously flawed. For BH, the energy is far too high compared to FCI or other methods, and for CH_4, the TCEPA curve lies significantly below FCI at intermediate and large internuclear distances. For Li_2, the TCEPA curve is somewhat close to the FCI curve, but again the energy is lower. The erratic behavior of the TCEPA results are not encouraging for this method by itself. Examining the curves from the TCEPA-CCSD hybrids discussed above, we see that they are generally improved over straight TCEPA. For BH and CH_4, TCEPA-CCSD(I) mimics TCEPA at short distances and significantly improves upon it at large distances. For Li_2, however, at large distances TCEPA-CCSD(I) ruins the fairly good TCEPA results. The more complete hybrid method TCEPA-CCSD(II) performs much better, yielding results quite close to CCSD (or to FCI) for BH or Li_2. For CH_4, the TCEPA-CCSD(II) curve is quite similar to that of CCSD or FCI, but it is shifted down to significantly lower energies, approximately matching the surprisingly low TCEPA energies at equilibrium.

Having discussed the qualitative features of the results, let us turn to a more quantitative assessment. The most important consideration is how parallel the approximate potential curves are to the exact FCI curves. This can be judged using the so-called non-parallelity error (NPE), which is defined as the difference between the largest error and the smallest error across a certain representative interval of interatomic separations. All the NPEs in this work were computed with respect to the FCI data. Table 1 presents the NPE's for eight molecules: H_2, BeH^+, BH, CH^+, Li_2, HF, H_2O. Based on the qualitative assessment above, it is no surprise that MP2 and CC2 exhibit very large NPE's over the intervals considered. Consistent with our previous work [41], NPE's for CCSD are modest for these reactions, ranging from less than 1 millihartree for Li_2 up to about 22 millihartree (13 kcal mol^{-1}) for HF and H_2O. While these errors are not acceptable for high-accuracy work, they may be acceptable in some applications, and they will be much lower for reactions in which bonds are made and broken simultaneously in the transition state [91].

Table 1: Non-parallelity errors in Hartree computed in the 6-31G* basis set. M-I, M-II, T-I, and T-II denote MP2-CCSD(I), MP2-CCSD(II), TCEPA-CCSD(I), and TCEPA-CCSD(II), respectively.

Molecule	Interval, Å	Step, Å	CCSD	MP2	CC2	M-I	M-II	TCEPA	T-I	T-II
H_2	0.5-3.9	0.1	0.0000	0.0728	0.0670	0.0049	0.0035	0.0060	0.0085	0.0012
BeH^+	0.6-4.1	0.1	0.0005	0.0456	0.0359	0.0070	0.0077	0.0290	0.0065	0.0054
BH	0.8-4.4	0.1	0.0084	0.0626	0.0447	0.0177	0.0075	0.0414	0.0224	0.0138
CH^+	0.6-4.0	0.1	0.0105	0.0616	0.0330	0.0338	0.0249	0.0778	0.0274	0.0125
Li_2	2.0-6.0	0.1	0.0002	0.0157	0.0123	0.0032	0.0011	0.0019	0.0041	0.0016
HF	0.8-3.8	0.2	0.0209	0.0624	0.0990	0.0265	0.0210	0.0302	0.0397	0.0338
CH_4	0.7-4.4	0.1-0.2	0.0172	0.0645	0.0536	0.0268	0.0169	0.0061	0.0324	0.0248
H_2O	0.7-4.0	0.1	0.0219	0.0837	0.0980	0.0264	0.0085	0.0464	0.0405	0.0343

Our first hybrid method based on MP2 amplitudes, MP2-CCSD(I), produces NPE's which are typically several times lower than those of MP2, but still somewhat larger than those of CCSD. MP2-CCSD(II) systematically improves the NPE's even further – it almost always works better than MP2-CCSD(I) and in four cases out of six (BH, HF, H_2O,and CH_4) rivals CCSD. Consistent with the erratic behavior of the TCEPA curves in Figs. 8-10, the NPE's of the TCEPA method are irregular. Except for the CH^+ molecule, they are lower than those of MP2, but this improvement is not predictable: sometimes TCEPA improves on MP2 by a factor of two or so (BeH^+, BH, HF, H_2O), and sometimes it even outperforms MP2-CCSD(II) and CCSD. The TCEPA-CCSD(I) and TCEPA-CCSD(II) methods are more systematic in this regard: the NPE's of TCEPA-CCSD(II) are always lower than those of TCEPA-CCSD(I) but they are still typically higher than those of MP2-CCSD(II). Figure 11 displays the NPE's averaged over the test cases considered here. Among the $\mathcal{O}(N^5)$ methods considered here, MP2-CCSD(II) performs best. It is remarkable that the average NPE of MP2-CCSD(II) is just as low as that of CCSD. The second best method is TCEPA-CCSD(II), which confirms our assumption that the inclusion of some additional amplitudes at the CCSD level should result in higher accuracy.

Although MP2 fails at large interatomic distances, it works well near the bottom of

Table 2: Spectroscopic constants of H_2, BeH^+, and BH computed using different methods in the 6-31G* basis set

Molecule	Method[a]	E_{min}	r_e	ω_e	$\omega_e x_e$	B_e	α_e	$\overline{D_e}$ (x10^{-4})
	MP2	-1.144141	0.7375	4533.58	126.1	61.502	3.0529	452.75
	CC2	-1.144174	0.7377	4527.85	126.7	61.466	3.0647	453.08
	MP2-CCSD(I)	-1.146218	0.7448	4367.81	143.2	60.305	3.4182	459.83
	MP2-CCSD(II)	-1.149402	0.7499	4297.58	140.1	59.483	3.3815	455.82
H_2	CCSD	-1.151698	0.7462	4367.09	141.7	60.080	3.3615	445.49
	TCEPA	-1.151003	0.7442	4403.38	141.8	60.404	3.3403	454.66
	TCEPA-CCSD(I)	-1.151508	0.7455	4384.19	140.1	60.199	3.3257	454.00
	TCEPA-CCSD(II)	-1.152484	0.7473	4354.30	141.1	59.910	3.3506	453.66
	FCI	-1.151698	0.7462	4367.09	141.7	60.080	3.3615	454.85
	MP2	-14.87313	1.3208	2280.90	34.3	10.660	0.2631	9.315
	CC2	-14.87338	1.3215	2275.49	34.5	10.649	0.2651	9.330
	MP2-CCSD(I)	-14.87368	1.3240	2251.12	37.8	10.609	0.2799	9.426
	MP2-CCSD(II)	-14.87656	1.3338	2176.70	42.6	10.454	0.3086	9.644
BeH^+	CCSD	-14.88154	1.3311	2193.06	40.4	10.496	0.2985	9.618
	TCEPA	-14.87941	1.3264	2269.61	39.1	10.571	0.2322	9.172
	TCEPA-CCSD(I)	-14.87980	1.3283	2252.92	43.2	10.541	0.2437	9.230
	TCEPA-CCSD(II)	-14.88146	1.3339	2214.01	44.3	10.451	0.2492	9.316
	FCI	-14.88159	1.3312	2192.20	40.4	10.495	0.2988	9.621
	MP2	-25.17587	1.2331	2451.40	47.6	12.007	0.3946	11.522
	CC2	-25.17634	1.2339	2443.36	48.3	11.993	0.3986	11.556
	MP2-CCSD(I)	-25.17660	1.2373	2399.63	55.1	11.926	0.4306	11.783
	MP2-CCSD(II)	-25.17780	1.2436	2336.48	57.7	11.805	0.4556	12.054
BH	CCSD	-25.20077	1.2443	2355.06	53.1	11.793	0.4281	11.183
	TCEPA	-25.20511	1.2371	2441.54	48.7	11.930	0.3828	11.393
	TCEPA-CCSD(I)	-25.20505	1.2370	2449.62	45.7	11.933	0.3723	11.327
	TCEPA-CCSD(II)	-25.20541	1.2393	2427.72	47.0	11.887	0.3797	11.400
	FCI	-25.20265	1.2448	2347.73	54.1	11.784	0.4333	11.874

the potential energy well. Therefore it is interesting to explore whether the new MP2-CCSD(II) method improves not only the behavior at large internuclear separations, but also the quality of results near equilibrium. If so, MP2-CCSD(II) might be preferable to MP2 not only for bond-breaking applications or cases where electronic near-degeneracies can become important, but also for routine computations of equilibrium molecular properties. Tables 2 and 3 present results for a number of spectroscopic properties, computed by fitting nine energy points evenly spaced by 0.005 Å about the equilibrium bond distance, r_e, to an eighth-order polynomial, $U(r)$. Each energy calculation was converged to at least 10^{-12} Hartrees and fitting errors are monitored to avoid numerical instabilities.

Table 3: Spectroscopic constants of CH^+, Li_2 and HF computed using different methods in the 6-31G* basis set

Molecule	Method[a]	E_{min}	r_e	ω_e	$\omega_e x_e$	B_e	α_e	$\overline{D_e}$ (x10^{-4})
CH^+	MP2	-37.96526	1.1195	3039.74	62.3	14.468	0.4876	13.111
	CC2	-37.96565	1.1199	3033.28	63.0	14.457	0.4911	13.136
	MP2-CCSD(I)	-37.97142	1.1143	3054.53	74.1	14.601	0.5326	13.346
	MP2-CCSD(II)	-37.97990	1.1160	3024.52	72.3	14.558	0.5385	13.492
	CCSD	-37.99427	1.1284	2930.87	68.5	14.240	0.5245	13.446
	TCEPA	-37.99884	1.1236	2955.43	43.5	14.362	0.5709	13.566
	TCEPA-CCSD(I)	-37.99592	1.1267	2915.13	37.2	14.282	0.5701	13.713
	TCEPA-CCSD(II)	-37.99712	1.1297	2882.33	21.0	14.208	0.5186	13.809
	FCI	-37.99628	1.1293	2919.43	69.5	14.218	0.5297	13.489
Li_2	MP2	-14.88685	2.7731	339.34	2.2	0.625	0.0050	0.085
	CC2	-14.88694	2.7753	337.67	2.2	0.624	0.0051	0.085
	MP2-CCSD(I)	-14.89129	2.7701	330.79	2.4	0.626	0.0055	0.090
	MP2-CCSD(II)	-14.89719	2.7387	334.71	2.8	0.641	0.0058	0.094
	CCSD	-14.89790	2.7254	340.09	2.8	0.647	0.0054	0.094
	TCEPA	-14.89943	2.7566	329.53	2.4	0.632	0.0054	0.094
	TCEPA-CCSD(I)	-14.89793	2.7381	339.06	2.3	0.641	0.0050	0.092
	TCEPA-CCSD(II)	-14.89932	2.7259	340.08	2.6	0.647	0.0054	0.094
	FCI	-14.89799	2.7249	339.96	2.7	0.647	0.0057	0.094
HF	MP2	-100.1842	0.9339	4040.83	83.5	20.196	0.7379	20.180
	CC2	-100.1851	0.9349	4019.61	84.8	20.153	0.7458	20.265
	MP2-CCSD(I)	-100.1845	0.9355	3989.97	91.4	20.125	0.7780	20.480
	MP2-CCSD(II)	-100.1856	0.9391	3922.40	90.0	19.974	0.7874	20.719
	CCSD	-100.1884	0.9342	4024.03	86.9	20.183	0.7543	20.309
	TCEPA	-100.2339	0.9446	3857.97	97.7	19.742	0.8143	20.678
	TCEPA-CCSD(I)	-100.2337	0.9433	3901.95	87.2	19.793	0.7672	20.373
	TCEPA-CCSD(II)	-100.2328	0.9420	3914.49	90.2	19.849	0.7798	20.413
	FCI	-100.1906	0.9355	3997.62	88.4	20.125	0.7634	20.401

[a] Core $1s$ electrons in carbon frozen

The familiar spectroscopic constants are computed by evaluating the zeroth to fourth-order derivatives of $U(r)$ at r_e and utilizing the following relations [92, 93]:

$$B_e \equiv \frac{h}{8\pi^2 \mu r_e^2}, \quad \omega_e \equiv \frac{1}{2\pi}\left(\frac{U''|_{r_e}}{\mu}\right)^{1/2}, \tag{71}$$

$$\omega_e x_e \equiv \frac{B_e^2\, r_e^4}{4h\omega_e^2}\left(\frac{10 B_e\, r_e^2\, \left(U'''|_{r_e}\right)^2}{3h\omega_e^2} - U^{\mathrm{iv}}|_{r_e}\right), \tag{72}$$

$$\alpha_e \equiv -\frac{2B_e^2}{\omega_e}\left(\frac{2B_e\, r_e^3\, U'''|_{r_e}}{h\omega_e^2} + 3\right), \overline{D}_e \equiv \frac{4B_e^3}{\omega_e^2}. \tag{73}$$

Here, μ is the reduced mass, B_e is the rotational constant, ω_e is the harmonic vibrational frequency, $\omega_e x_e$ is the anharmonicity constant, α_e is the vibration-rotation coupling constant, and $\overline{D}_e$ is the centrifugal distortion constant.

Regarding the total energies at equilibrium, $E_{\min}$, note that those of the TCEPA-based

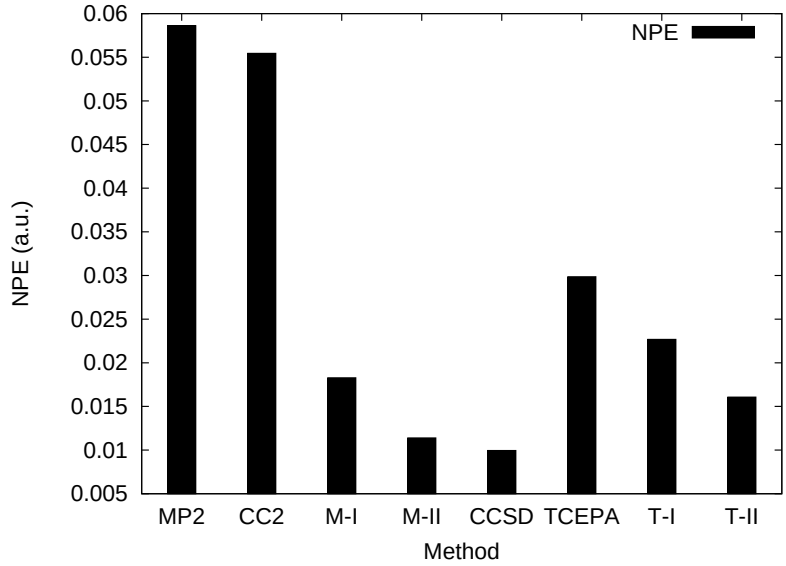


Figure 11: The average non-parallelity errors (NPE) in 6-31G* basis set relative to FCI.

methods are often significantly lower than those of FCI (as mentioned above). The MP2-based hybrid methods as well as CC2 tend to act like MP2 itself in their prediction of $E_{\min}$. Considering the equlilibrium bond length r_e, MP2 systematically underestimates this parameter (except for Li_2), while TCEPA behaves irregularly. Figure 12 shows the root mean square (RMS) errors for the spectroscopic constants. After CCSD, the lowest RMS values of r_e belong to TCEPA-CCSD(II), TCEPA-CCSD(I) and MP2-CCSD(II). The individual equilibrium distances produced by the MP2-CCSD methods are typically smaller than those produced by the TCEPA-CCSD methods. The constants α_e and $\omega_e x_e$ depend on the third and fourth derivatives, respectively, of the potential and so are sensitive to the shape of the potential. MP2-CCSD(II) is the best performer out of hybrid methods for these constants, while the TCEPA-based hybrid methods are inferior to the MP2-based hybrids for these characteristics.

Somewhat unexpectedly, the errors of MP2-CCSD(II) for the α_e, ω_e, and $\overline{D}_e$ constants frequently have the sign different from those of all other methods. The centrifugal distortion constant $\overline{D}_e$ is estimated with similar quality by MP2-CCSD(II) and the TCEPA-based

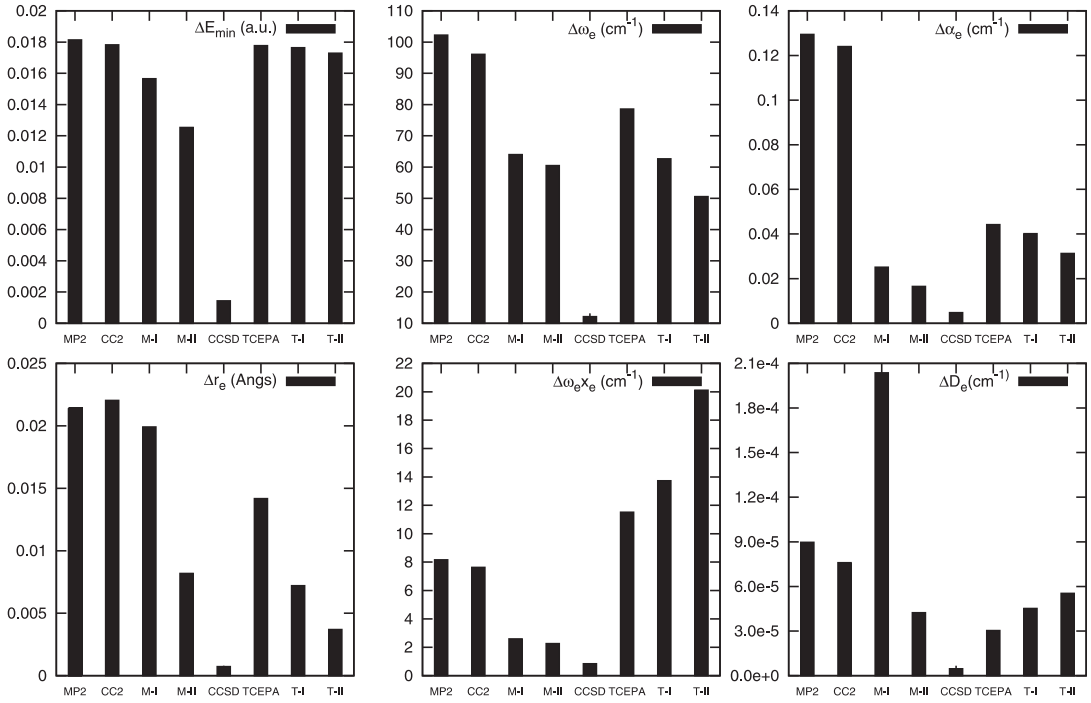


Figure 12: The root mean square (RMS) errors of various spectroscopic constants in 6-31G* basis set relative to FCI. M-I, M-II, T-I, and T-II denote MP2-CCSD(I), MP2-CCSD(II), TCEPA-CCSD(I), and TCEPA-CCSD(II), respectively.

methods, all of which perform better than MP2 or CC2. The larger RMS error observed for MP2-CCSD(I) is due almost entirely to a single poor result for the H_2 molecule. RMS errors for B_e are omitted from Figure 12 because this characteristic is proportional to r_e^{-2} and its errors are tied to r_e errors. Judging from the magnitude of the RMS errors of the spectroscopic constants, we conclude that MP2-CCSD(II) is the most consistent among the $\mathcal{O}(N^5)$ methods near equilibrium. The CC2 method shows only a slight improvement over MP2, while the TCEPA-based methods are apparently the worst performers.

4.5 Conclusions

In this work we have employed hybrid methodology to construct several new methods referred to as MP2-CCSD(II), TCEPA-CCSD(I), and TCEPA-CCSD(II) anticipating to find an $\mathcal{O}(N^5)$ scheme that improves upon the performance of the previous, MP2-CCSD(I),

method. The computation of the NPE's and several spectroscopic constants for a number of simple molecules has shown that MP2-CCSD(II) works noticeably better than MP2 and sometimes rivals even CCSD, which scales as $\mathcal{O}(N^6)$. The average NPE error of MP2-CCSD(II) is not worse than that of CCSD. The simplicity of formulation, the inexpensiveness and the accuracy of the MP2-CCSD(II) method express the hope that it might be used instead of MP2 in many situations where the latter is currently applied.

A few limitations of MP2-CCSD(II) (equally applicable to our other hybrid methods) must be mentioned, however. First, we do not expect it to exhibit an impressive performance in cases where CCSD itself should fail. Such cases may include breaking multiple bonds or other cases of strong electronic near-degeneracies. A more sophisticated hybrid scheme may be desirable to deal with these issues. For example, the inclusion of higher than double excitations or accounting for the multireference character of the ground state may be needed. Indeed, work is in progress on such schemes as MP2-CCSDTQ and MP2-MCSCF which will be more suitable to conform to these requirements. Second, as all the methods based on the active space partitionings, MP2-CCSD(II) obliges the researcher to select a proper active space. We believe than in many cases the minimal dimension of the active space (i.e., only σ and σ^* for single bonds) should be satisfactory, but sometimes slightly larger active spaces may be required. Such complications arise when there are orbitals whose energies are very close to those of σ and σ^*, or when σ and σ^* are of different character at different interatomic separations.

An attractive application for the new MP2-CCSD(II) method would be to systems for which CCSD performs well but MP2 misbehaves. In a separate study[91], where we investigate radical hydrogen abstraction reactions, the enthalpies produced by CCSD with non-iterative triples, CCSD(T), are relatively insensitive to the choice of the reference function, whether UHF or ROHF. However, the enthalpies computed with MP2 depend crucially on the choice of the reference. Additionally, the ROHF-based MP2 method produces some

unacceptable artifacts whereas the UHF-based MP2 method suffers from serious spin contamination. We believe that the MP2-CCSD(II) method would alleviate such problems of MP2 if applied to these systems and plan to explore this in future work.

The material of this chapter was submitted as a paper to the *Journal of Chemical Physics* [94].

CHAPTER V

THE ELECTRON AND NUCLEAR ORBITAL MODEL: CURRENT CHALLENGES AND FUTURE PROSPECTS

5.1 Introduction

The Born-Oppenheimer (BO) approximation [29] defines the electronic wave function and the electronic spectrum as having a parametric dependance on the clamped nuclei's locations in space. Although it is suitable for the description of the great majority of cases, recently there has been a lot of development and implementation of non-BO theoretical schemes [95, 96, 97, 98] and corrections to the BO surfaces [99, 100, 101, 102, 103, 104, 105]. Time-dependent approaches to dynamics of electrons and nuclei in molecules are summarized in a comprehensive review [106] and further explored in [107]. The development of these *non-adiabatic* approaches is driven partly by the ever increasing accuracy of traditional BO *ab initio* quantum chemical methods such that non-BO corrections become one of the residual sources of error in theoretical spectra of small molecules. in addition to offering a more rigorous description of molecular structure of interest to spectroscopists, non-adiabatic wave function ansatze allow a natural description of systems with strongly interacting or crossing electronic BO surfaces and of effects of quantum nuclei, such as breaking and making of bonds to hydrogen and other light nuclei through tunnelling processes. Certainly, more general non-BO approaches are highly desirable. In this work we investigate one such recently proposed approach which is attracting considerable attention.

The mean-field description of the electrons, which move in the clamped nuclei's external electric field [6, 7, 9, 10], underprops most of the models currently used in quantum chemistry. Now it seems straightforward to construct a more general, non-Born-Oppenheimer model, which describes both the electrons and the nuclei in a mean-field approximation. In such a model, the nuclei can be regarded as particles that are distinguished from the electrons in the following aspects: (i) they have different masses; (ii) they possess positive electric charges; (iii) they do not exhibit exchange effects with respect to the electrons; (iv) they may have different spin statistics.

Such a self-consisted field (SCF) formalism provides the natural start for a molecular structure theory (as opposed to electronic structure theory) and supplies a reference electronic-nuclear wave function for all the subsequent hierarchies of models which include interparticle correlation – the analogs of many-body perturbation theory (MBPT), configuration interaction (CI), coupled-cluster (CC) and multiconfigurational self-consisted field (MCSCF) methods. Let us consider the simple case when there is only one type of fermion nuclei of spin $1/2$ present in the molecule. Since the extension to several types of nuclei (including bosons) is trivial, we will adhere to this simple case throughout the chapter. The wave function ansatz upon which the molecular structure theory may be easily built has the following form:

$$\Psi \equiv \Psi(r_1, r_2, ..., r_N,\ R_1, R_2, ..., R_M) = \frac{1}{\sqrt{N!\,M!}} \begin{vmatrix} \varphi_1(r_1) & \varphi_2(r_1) & ... & \varphi_N(r_1) \\ \varphi_1(r_2) & \varphi_2(r_2) & ... & \varphi_N(r_2) \\ ... & ... & ... & ... \\ \varphi_1(r_N) & \varphi_2(r_N) & ... & \varphi_N(r_N) \end{vmatrix} \times \begin{vmatrix} \Phi_1(R_1) & \Phi_2(R_1) & ... & \Phi_M(R_1) \\ \Phi_1(R_2) & \Phi_2(R_2) & ... & \Phi_M(R_2) \\ ... & ... & ... & ... \\ \Phi_1(R_M) & \Phi_2(R_M) & ... & \Phi_M(R_M) \end{vmatrix}, \tag{74}$$

where $\varphi_q(r)$ and $\Phi_q(R)$ are electronic and nuclear spin-orbitals, respectively, which depend on electronic r and nuclear R coordinates. We call models based on this ansatz electronic and nuclear molecular orbitals (ENMO) methods.

The first attempt to apply ansatz (74) to the calculations of the properties of molecules was probably made by Thomas [108, 109]. He presented the basic mean-field calculations on methane, ammonia and water in which the hydrogen atoms were treated quantum-mechanically and the heavy atoms classically. It seems that until recently not much attention has been paid to further development of this model. However, over the last few years several works appeared in which ansatz (74) was used. Tachikawa [110] and co-workers presented the mean field equations in the usual quantum-chemical notation and applied the method to the $[e^+;\mathrm{F}^-]$, HF and DF systems using variationally optimized uncontracted Gaussian basis sets. They called the mean-field method based on ansatz (74) dynamically extended MO (DEMO). Later Tachikawa and Osamura [111] performed first mean-field ENMO calculations on different isotopomers of H_2 as well of those of LiH in which all the nuclei were treated quantum-mechanically. These works led to a series of applications particularly on the isotope effects in hydrogen-containing molecules [112, 113, 114]. The valence bond [115] and CIS (configuration interaction with single excitations only) [116] extensions of ENMO were also applied to the isotopomers of H_2 and the latter to H_3^+. Recently, Hammes-Schiffer and co-workers [117] derived the corresponding CI and MCSCF equations under acronym NEO, and Tachikawa [118] performed some full configuration interaction (FCI) calculations on H_2 and its isotopomers. Later, Nakai and Sodeyama [119] developed MBPT and CC versions of the method.

A theory that describes all the electrons and the nuclei in the molecule simultaneously discards the notion of the potential energy surface (see an interesting discussion in Ref. [120]). This radical step allows, albeit in principle, 'geometry' optimization in one non-iterative calculation as well as the calculation of dynamical processes, for example, the tunnelling of light atoms. However, a number of hidden difficulties are inherent in this

simple approach. Apart from those difficulties that we discuss in great detail later in the chapter, we see at least two which make it hard for constructions of type (74) to serve as a universal ansatz for non-adiabatic quantum chemistry. First, wave function (74) aspires to describe the molecule, or the molecular system, in all its possible classical geometries. The existence of such dissociative pathways places corresponding full molecular wave functions outside of the Hilbert space. Proper description od such situations would be similar to that in scattering calculations, that is technically rather more complicated than that of the usual bound BO electronic structure theory. Even for closed channels, knowing how drastically different the BO wave function may be, say, for the equilibrium positions of the nuclei and near dissociative limit, we cannot but assume that in order to guarantee the sufficient flexibility of (74), the number of the orbitals in it should be enormous, possibly orders of magnitude larger than currently considered tractable. Secondly, there is no explicit dependence on the interparticle distance in ansatz (74). Thus, efficient description of the Kato cusp [121] is not possible and may lead to slow asymptotic convergence of CI-type based on (74) as a reference.

At present there is no common agreement about several important issues in the ENMO approach. In some works [109, 116, 115, 118] it was considered important to separate the center of mass (CM) motion from the Hamiltonian whereas in others [111, 117, 119] a CM transformation has not been performed for the molecules in which all the nuclei were described quantum-mechanically. These two different treatments of the CM issue lead to clearly different qualities of results, such as total nonadiabatic energies, which, however, are not reported in some works [117, 119], meaning it hard to interpret the other presented numbers.

In the present work we clarify this and a number of other questions by presenting the first fundamental study examining the convergence of ENMO approaches with respect to basis set and correlation treatment. In comparison to the hydrogen atom, for which analytic solutions exist, we find that spurious, non-physical states clutter the spectrum. By fixing

at least one nucleus in the molecule, the translational invariance is broken and the spurious states may be removed. However, we show for the first time that this hybrid approach, in which some nuclei are treated classically and some quantum-mechanically, does not provide very accurate results for vibrations involving both quantum and classical nuclei. We believe that the successful resolution to the problem raised in this work will be critical to future advances.

5.2 Electronic and nuclear molecular orbital (ENMO) method

The general theory of the ENMO method is described in a few articles (see, for example, Ref. [117]). Here we would like to briefly mention the gist of the approach and concentrate on some important aspects.

It should be noted from the very beginning that the ENMO theory is almost covariant to the MO theory. Under covariance of two models we mean the identity of the general form of their equations. Points (ii) and (iii) defined in the introduction break the full covariance of the ENMO to the MO. If the nuclear orbitals $\Phi_1, \Phi_2, ...\Phi_M$ are denoted as $\varphi_{N+1}, \varphi_{N+2}, ...\varphi_{N+M}$ and the nuclear coordinates $R_1, R_2, ...R_M$ as $R_{N+1}, R_{N+2}, ...R_{N+M}$ the wave function (74) will take form of the ordinary Slater determinant

$$\Psi(r_1, ..., r_K) = \frac{1}{\sqrt{N!\,M!}} \begin{vmatrix} \varphi_1(r_1) & \varphi_2(r_1) & ... & \varphi_K(r_1) \\ \varphi_1(r_2) & \varphi_2(r_2) & ... & \varphi_K(r_2) \\ ... & ... & ... & ... \\ \varphi_1(r_K) & \varphi_2(r_K) & ... & \varphi_K(r_K) \end{vmatrix} \tag{75}$$

where $K = N + M$ and

$$\varphi_1(r_L),\ \varphi_2(r_L), ..., \varphi_L(r_L) = 0, \quad L = N+1, N+2, ..., K \tag{76}$$

$$\varphi_L(r_{N+1}),\ \varphi_L(r_{N+2}), ..., \varphi_L(r_K) = 0, \quad L = 1, 2, ..., N. \tag{77}$$

If the normalization factor $1/(\sqrt{N!\,M!})$ were equal to $1/(\sqrt{K!})$ and the conditions (76) (77) were lifted then the ENMO and MO would be fully covariant. However, even these

constraints leave a good deal of covariance, which is very useful for constructing further models. It is convenient for the future discussion to define the Hamiltonian H as flexible with respect to the CM transformation. Therefore we write H as:

$$H = T^e + T^n + V^{ee} + V^{en} + V^{nn} - \lambda\, T_{CM}, \tag{78}$$

where T's are the kinetic energy operators and V's are Coulomb interaction operators. The symbols e, n, CM stand for electrons, nuclei and center of mass, respectively. The constant λ takes on the two values - either 1 or 0, depending on whether we perform the CM transformation or not. The kinetic energy of the CM operator reads [122]

$$T_{CM} = -\frac{1}{2M_T}(\sum_{i}^{K} \nabla_i^2 + \sum_{i\neq j}^{K} \nabla_i \nabla_j) \tag{79}$$

with M_T being the mass of the molecule. T_{CM} may be used to redefine T^e, T^n and the Coulomb interaction operators to cast the Hamiltonian H into the covariant form

$$H = T'^e + T'^n + V'^{ee} + V'^{en} + V'^{nn} \tag{80}$$

where

$$T'^e = -\sum_{i}^{N} \frac{1}{2}(\frac{1}{m_e} - \frac{\lambda}{M_T})\nabla_i^2, \tag{81}$$

$$V'^{ee} = \sum_{i<j}^{N}(\frac{1}{r_{ij}} + \frac{\lambda}{M_T}\nabla_i \nabla_j) \tag{82}$$

and the analogous expressions for T^n, $V^{(en)}$, and $V^{(nn)}$. The variation of the energy

$$E = \langle\Psi|H|\Psi\rangle / \langle\Psi|\Psi\rangle \tag{83}$$

gives the following set of coupled differential equations:

$$\left[T'^e + \sum_{k=1}^{K} J_k - \sum_{k=1}^{N} K_k\right]\varphi_i = \varepsilon_i \varphi_i \tag{84}$$

for $i = 1, 2, ..., N$ (electronic equations) and

$$\left[T'^n + \sum_{k=1}^{K} J_k - \sum_{k=N+1}^{K} K_k\right]\varphi_i = \varepsilon_i \varphi_i \tag{85}$$

for $i = N+1, N+2, ..., K$ (nuclear equations). The operators J_k and K_k are equivalent to the Coulomb and exchange operators of the Hartree-Fock (HF) theory. When integrating spins out and leaving only spatial functions in φ_i's, we arrive at four different mean-field theories: RHF/RHF, RHF/UHF, UHF/RHF and UHF/UHF where the abbreviation before the slash refers to the electronic part and after the slash refers to the nuclear part. It is not surprising that the unrestricted nuclear orbitals are more suitable for the description of molecular structure than the restricted ones. The nuclear orbitals are so localized that each nucleus has to be characterized by its own spatial wave function, which corresponds to the UHF method. Our experience shows that in all practical cases the RHF/UHF and UHF/UHF theories give much lower energy than their .../RHF analogs. Therefore, in what follows we will outline only the UHF/UHF version, it being more general than RHF/UHF one.

The total energy of ENMO-UHF/UHF is given by the following expression:

$$E = \frac{1}{2}\mathrm{Tr}(G^{e\alpha}P^{e\alpha} + G^{e\beta}P^{e\beta} + G^{n\alpha}P^{n\alpha} + G^{n\beta}P^{n\beta}) \tag{86}$$

where $G = T' + F$ and F, P are the Fock and density matrices, respectively. The matrix elements of Fock matrix in an MO basis are (the case of only $F^{e\alpha}$ and $P^{e\alpha}$ is shown; all the other matrix elements may be written analogously):

$$F_{ij}^{e\alpha} = T_{ij}^{\prime e\alpha} + \sum_{k}^{N^{e\alpha}}[(i_\alpha^e j_\alpha^e|k_\alpha^e k_\alpha^e) - (i_\alpha^e k_\alpha^e|j_\alpha^e k_\alpha^e)] +$$

$$\sum_{k}^{N^{e\beta}}(i_\alpha^e k_\alpha^e|j_\beta^e k_\beta^e) - \sum_{k}^{N^{n\alpha}}(i_\alpha^e j_\alpha^e|k_\alpha^n k_\alpha^n) - \sum_{k}^{N^{n\beta}}(i_\alpha^e j_\alpha^e|k_\beta^n k_\beta^n). \tag{87}$$

Here $N^{e\alpha}$ designates the number of alpha electrons and $T_{ij}^{\prime e\alpha}$ are the matrix elements of the primed kinetic energy operator. The two-particle integrals like $(i_\alpha^e k_\beta^e|j_\alpha^e k_\beta^e)$ in (87) are defined for the primed Coulomb operators (such as (82)) and absorb the particles' electric charges. Further, following the idea of Roothaan [9, 10], each molecular orbital $|i_\alpha^e\rangle$ is expressed as a linear combination of basis orbitals with coefficients $C^{e\alpha}$. The coefficients

of such a decomposition must satisfy the set of Pople-Nesbet-like equations:

$$F^{e\alpha}C^{e\alpha} = S^e C^{e\alpha}\varepsilon^{e\alpha} \tag{88}$$

$$F^{e\beta}C^{e\beta} = S^e C^{e\beta}\varepsilon^{e\beta} \tag{89}$$

$$F^{n\alpha}C^{n\alpha} = S^n C^{n\alpha}\varepsilon^{n\alpha} \tag{90}$$

$$F^{n\beta}C^{n\beta} = S^n C^{n\beta}\varepsilon^{n\beta}. \tag{91}$$

with S^e and S^n being the electronic and nuclear overlap matrices.

5.3 The center-of-mass separation

Expressions (88)-(91) were derived from (83) with the help of the variation principle. Yet the application of the variation principle is justified only when the Hamiltonian possesses bound eigenstates [123]. Since the Hamiltonian (78) with $\lambda = 0$ necessarily contains the CM motion, all of its eigenstates must include a continuous CM part. Therefore, if we assume that the system in question possesses bound states [124], we have to put $\lambda = 1$ in (78) for the variation principle to be safely applied to its expectation value (83). Now that the Hamiltonian does not include the CM motion, there emerges another difficulty. The CM motion-free Hamiltonian is translationally invariant, yet the wave function (74) is not invariant due to the space-fixed one-particle basis set. A proper approach to solving such Schrödinger equation is to utilize 3K-3 internal coordinates directly [**?**]. Of course, the wave function ansatz in such approach has no resemblance to ansatz (74) composed of the product of *one-particle* functions. To get rid of the spurious states, we could re-write and solve the problem in the internal, translationally-invariant coordinates. The implementation of this option is rather complicated. In doing so, we would lose the simplicity and flexibility of the ENMO approach. Instead, we could choose to fix one of the particles in the molecule (preferably, the heaviest nucleus). Although this option has a number of theoretical disadvantages in comparison with the previous one, practically, it seems much

more appealing to us since, formally, its acceptance does not modify any formulas of the ENMO approach. In practice, to fix a nucleus, one just has to place one s-function with an 'infinite' nuclear exponent α on it together with making its mass m 'infinite' as well, but keeping the ratio $\alpha^2/m \to 0$. Then one does not even have to calculate the matrix elements of the T_{CM} operator since it is equal to zero. In this work, however, we would also like to demonstrate what consequences stem from not accepting either of the above-mentioned options, as it was, for instance, in Refs. [110, 112, 113, 111, 115, 119].

Further, it is noteworthy that the correct eigenfunction of Hamiltonian (78) must be a product of the internal and center-of-mass components. Even though trial function (74) is not separable in general case, the variation of (83), as we shall demonstrate below, seems to work in a pragmatic sense. Still, the non-separability manifests itself in poor convergence of the self-consistent iterations and their frequent arrival at higher-energy states. To converge the ENMO-SCF energy to a high precision, such as 10^{-8} Hartree, it sometimes required hundreds of iterations, quite irrespective of the structure of the guess matrices and the number of the guess Fock matrices in the direct inversion of the iterative subspace (DIIS [125, 126]) acceleration algorithm.

5.4 The basis sets

The basis sets used in this work were constructed from uncontracted Cartesian Gaussians. Since the full molecular Hamiltonian is invariant with respect to arbitrary relations of the space, exact wave functions must be eigenfunctions of the orbital angular momentum operators ($\hat{L}^2$ and $\hat{L}_z$). It is clear that ansatz (74) describes eigenfunctions of those operators if all Gaussian functions are placed at the same point. Thus, for the hydrogen atom we placed electronic and nuclear functions at the origin. In molecules, such placement of basis functions endangers efficient description of electronic, and, especially, nuclear orbitals. Hence, out of practicality one has to place nuclear basis functions near their most probable classical positions. Corresponding electronic wave functions are placed at each such 'nuclear

center' so as to minimize errors due to the lack of electron cusp in (74).

The basis sets are denoted as [A:B] where A and B refer to electronic and nuclear bases, respectively. Because of the large number of basis functions required for our calculations, almost all of our basis sets were even-tempered. The exponents were not optimized tightly, according to the general recommendation [127] that it is more advantageous to add extra functions than to optimize the exponents exhaustively. To minimize the number of optimized variables, we assumed that for atoms the ratio of electronic and nuclear exponents should be equal to the ratio of their masses. Interestingly, it is easy to show that for the case of hydrogen atom with one electronic and one nuclear basis function keeping the ratio of exponents equal to the inverse mass, leads to the perfect factorizability of the wave function into the CM and the internal components. For molecules, nuclear exponents must be kept moderate in order to describe the internuclear vibrations correctly.

5.5 The MO limit

Now we would like to comment on the possible comparison of the ENMO energies with those of the ordinary MO method. Tachikawa and Osamura reported [111] their best mean-field-ENMO energy for the hydrogen molecule equal to -1.052641 Hartrees, obtained with a [5s:1s1p1d] basis set without the CM separation, and compared this value to the corresponding MO result, -1.13217 Hartrees. They explained the discrepancy by arguing that 'the nuclear kinetic energies are not included in the conventional MO method'. We believe that the high energy of -1.052641 Hartree was the result of the non-subtraction of CM kinetic energy from the Hamiltonian, apart from the poor quality of the basis set.

The CM kinetic energy operator is positive-definite and hence the lowest expectation value of the Hamiltonian $H_{\lambda=1}$ is lower or equal to the lowest expectation value of the Hamiltonian $H_{\lambda=0}$. We found that the CM separation in ENMO results in a substantial lowering of the energy in all practical cases. For example, for the hydrogen molecule with the Tachikawa and Osamura basis set [5s:1s1p1d] $H_{\lambda=1}$ gives a significant improvement of

-1.074653 Hartrees compared to the $H_{\lambda=0}$ result.

Instead, the ENMO energy should have been compared to the non-adiabatic energy -1.164024 Hartree [128]. This raises an interesting question of under what limiting conditions the ENMO model becomes the MO model. It is easy to demonstrate that if there is only one nuclear orbital on the nucleus in the hydrogen atom and if the nuclear exponent is increased simultaneously with the nuclear mass so that the expression $\alpha^2/m \rightarrow 0$ is satisfied, the ENMO energy converges to the MO limit (see Figure 13). If there are several nuclei in the molecule the MO limit is attained by fixing all the nuclei by the same procedure.

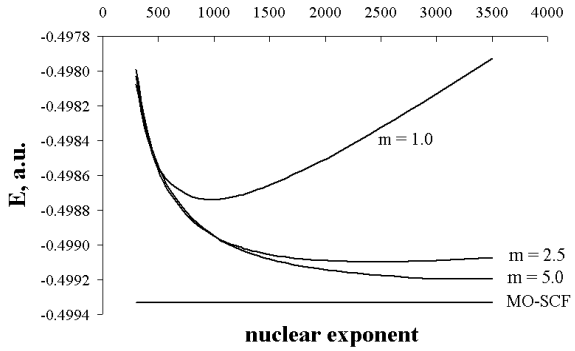


Figure 13: The transition of the ENMO-SCF energy to the MO-SCF limit for the hydrogen atom in [aug-cc-pvDZ:1s] basis set. Each curve is generated for different mass of the nucleus m which is given in the units of proton mass. In the limit of an infinite exponent and infinite nuclear mass, under the condition $\alpha^2/m \rightarrow 0$, the ENMO-SCF energy reaches down to the MO-SCF energy.

5.6 Interparticle correlation: ENMO-MBPT and ENMO-CI

It is possible to improve the mean-field results by taking into account electronic-electronic, electronic-nuclear and nuclear-nuclear correlation. The correlation in ENMO model may be handled exactly in the same way as it is done in the BO-MO model. The formulas for ENMO-MPBT2 and the ENMO-CISD are straightforward to derive and were reported in

Refs. [119] and [117], respectively.

For the hydrogen atom the ENMO-CISD should provide the exact answer in the limit of an infinite basis set because the ENMO-CISD is equivalent to the ENMO-FCI in this case. The spectrum of the hydrogen atom produced by the ENMO-CISD may be compared to the exact analytical result and that is why this test becomes crucial for understanding the specifics of the ENMO description of non-adiabatic problems.

Table 5.6 presents the non-adiabatic ground-state energies of the hydrogen atom calculated by the ENMO-MPBT2 and the ENMO-CISD with a series of basis sets of increasing quality, where the CM kinetic energy has been removed. The parameters section should be interpreted as follows. The first number in every column shows the largest exponent taken for the corresponding orbital type of the electronic basis set. The second number represents the factor by which every subsequent electronic exponent is divided in order to obtain the chain of even-tempered exponents. All the nuclear exponents, except for the d-functions, were taken proportional to the electronic ones with the factor of proportionality 1836.1558. For nuclear d-functions, the analogous parameters are given after the colon. The stars indicate basis sets that are not completely even-tempered. To construct a singly-starred basis one needs to proceed according to the above guidelines, but change the lowest electronic p-exponent to 0.13. A doubly-starred basis is constructed analogously to the singly-starred one, but the lowest electronic p-exponent is 0.06, and the second-lowest is 0.13. The basis set [12s7p3d1f:8s6p3d]** contains an f-function (not mentioned in the table explicitly, for the compactness of representation) with parameters $700.0\ 1.00$. The table shows shows that the ENMO-CISD performs much better than the ENMO-MBPT2 but even the ENMO-FCI convergence to the exact non-adiabatic energy with respect to the basis set is apparently rather slow.

For comparison, the traditional MO-SCF with the electronic part of the [12s7p3d1f:8s6p3d]** basis set gives the 0.0001% error of ground-state energy whereas the error of ENMO-FCI

Table 4: The convergence of the nonadiabatic energy of the H atom as a function of basis set. The analytical result is -0.49972784 Hartrees. All energies are given in Hartrees. For the explanation of the basis set parameters see the text.

Basis	Basis set parameters			ENMO-UMP2	ENMO-CISD
set	s	p	d	energy	energy
[10s:7s]	180.0 2.6000			−0.49934777	−0.49934847
[11s:7s]	180.0 2.6000			−0.49934816	−0.49934882
[12s:8s]	180.0 2.2794			−0.49935094	−0.49935170
[12s2p:8s2p]	180.0 2.2794	3.0 11.00		−0.49937520	−0.49952536
[12s3p:8s3p]	180.0 2.2794	13.0 7.00		−0.49941720	−0.49959414
[12s4p:8s4p]	180.0 2.2794	50.0 5.75		−0.49944467	−0.49962935
[12s5p:8s5p]	180.0 2.2794	70.0 4.00		−0.49945476	−0.49965024
[12s6p:8s6p]	180.0 2.2794	95.0 3.20		−0.49946026	−0.49966233
[12s6p:8s6p]*	180.0 2.2794	420.0 3.8		−0.49946749	−0.49971395
[12s7p:8s6p]**	180.0 2.2794	420.0 3.8		−0.49946806	−0.49971478
[12s7p1d:8s6p1d]**	180.0 2.2794	420.0 3.8	400.00 1.0 : 1010900.0 1.0	−0.49947002	−0.49971748
[12s7p3d:8s6p1d]**	180.0 2.2794	420.0 3.8	1000.0 2.0 : 1010900.0 1.8	−0.49946981	−0.49971779
[12s7p3d:8s6p3d]**	180.0 2.2794	420.0 3.8	1000.0 2.0 : 1010900.0 1.8	−0.49947034	−0.49971780
[12s7p3d1f:8s6p3d]**	180.0 2.2794	420.0 3.8	1000.0 2.0 : 1010900.0 1.8	−0.49947073	−0.49971833

with the [12s7p3d1f:8s6p3d]** basis set is 0.002%. The slow convergence is due to the inability of trial wave functions composed of products of single-particle functions to describe the particle-particle cusp [121]. Unfortunately, it is hard to judge whether such a poor convergence of energy to the exact result was also observed in the ENMO-FCI study of the hydrogen molecule by Tachikawa [118] because the best basis set employed in that work was too small to draw any definite conclusions. The non-adiabatic ansatz of Adamowicz and co-workers [129] includes explicit dependence on the interparticle distances in the wave function due to the use of Gaussian geminals. It is then not surprising that despite the aforementioned difficulties of constructing fully non-adiabatic wave functions, it is possible to obtain results of high accuracy with their method. Indeed, it takes only 14 explicitly correlated functions to reach the energy -0.499724 Hartree for the hydrogen atom [95, 96, 97, 98, 129]. It takes only several hundred of explicitly correlated functions to obtain a result of comparable accuracy for the simplest molecular systems such as H_2^+ or H_2, a rather efficient expansion indeed, when compared to the number of determinants required to reach a similar accuracy in traditional electronic structure calculations.

However, the poor convergence of the energy is not the most serious drawback of the ENMO model. In Figure 14 the spectrum of the hydrogen atom generated by the ENMO-FCI method in our biggest basis set is presented.

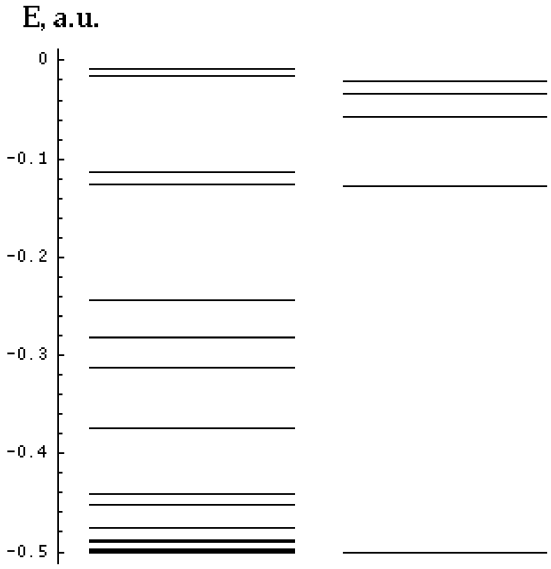


Figure 14: On the left: The spectrum of the hydrogen atom produced by the ENMO-CISD with the [12s7p3d1f:8s6p3d] basis set. On the right: The analytical spectrum of the hydrogen atom.

It is clear that the ENMO-FCI, apart from the first few genuine excited states, yields a lot of spurious states which are not predicted by the analytical solution. Moreover, some of the spurious states are three- and five-fold degenerate. These results demonstrate that the simple ENMO approach is unsuitable for the computation of excited states. The incorrect spectrum of the Hamiltonian is clearly the result of solving the problem in a larger coordinate space than is required. IN nuclear structure theory, the spurious states problem is a notorious feature also, due to the translational invariance of the nuclear Hamiltonian. In that case, it is possible to circumvent the problem by placing the system into a harmonic well and minimizing the well contribution to the internal spectrum of the system [130, 131]. This is only achievable by virtue of using the same mass for both types of nucleons and the choice of harmonic oscillator eigenfunctions for the one-particles basis. Whether a similar

approach is feasible within the ENMO framework is not obvious to us.

The presence of spurious states is a major obstacle which does not allow the ENMO model to serve as the full-fledged generalization of the conventional MO model. However, 'fixing' the heaviest nucleus in a molecule and treating all the rest of the particles quantum-mechanically breaks the translational invariance of the Hamiltonian and eliminates the spurious states from the spectrum. For atoms, that approach is equivalent to the MO model. For molecules, the heavier the fixed nucleus, the better non-adiabatic description this approach gives. One can also freeze several, or all but one nucleus in the molecule. In the rest of the chapter we will refer to this improved ENMO model as the 'hybrid' ENMO model, to the frozen nuclei as 'classical' and to the rest of the nuclei as 'quantum'.

5.7 Vibrational spectra of diatomic molecules

One of the features of the hybrid ENMO approach lies in its ability to predict vibrational states that correspond to the vibrations of the 'quantum' nuclei simultaneously with the electronic states. Now we are chiefly interested in answering the question of whether the hybrid ENMO model can predict the energies of the vibrational states accurately, or, alternatively, what magnitude of error is introduced as we freeze one of the nuclei. To keep the discussion unencumbered by extra details, we will consider as the objects of our study the simple diatomic hydrides in which the proton is treated classically and the other nucleus quantum-mechanically. We wish to estimate the performance of the ENMO-CI in this simplest case. The harmonic frequency of the mode in question is

$$\omega = \sqrt{k/\mu}, \tag{92}$$

where k is the force constant and μ is a reduced mass. Keeping in mind that k is defined primarily by the electromagnetic forces that act between the atoms in the molecule, we assume that it does not depend on the mass of the atoms. Then it is possible to calculate ω as the function of the reduced mass μ. The harmonic vibrational frequency of the diatomic

Table 5: The estimate of the hybrid ENMO-CI performance for the diatomic hydrides. The error is calculated as $|(\omega_\infty - \omega)/\omega|$. All the values refer to the ground electronic states.

Molecule	ω, cm^{-1}	k, a.u.	ω_∞, cm^{-1}	Error, %
H_2	4401.23	$2.03 \cdot 10^{-4}$	3112.14	29.3
^{7}LiH	1405.65	$6.61 \cdot 10^{-5}$	1314.41	6.49
$H^{19}F$	4138.33	$3.40 \cdot 10^{-4}$	4032.74	2.55

hydride with one atom frozen becomes

$$\omega_\infty = \sqrt{k/m_p}, \tag{93}$$

where m_p is the proton mass. The constant k may be calculated from the experimental reduced mass and vibrational frequency. The estimate ω_∞ would be the best result the hybrid ENMO-CI would achieve if it did not automatically include anharmonic corrections. However, for most molecules, these corrections are rather small and ω_∞ should be a good rough estimate for the value of ω the ENMO-CI should produce (theoretically, at the level of the full configuration interaction and the infinite basis set limit). Table 5 shows ω_∞'s for the hydrides studied in this work. Table 2. The estimate of the hybrid ENMO-CI performance for the diatomic hydrides. The error is calculated as $|(\omega_\infty - \omega)/\omega|$. All the values refer to the ground electronic states.

The table demonstrates that even for the diatomic hydrides in which the heavy nucleus is many times heavier than the proton (as in $H^{19}F$), the expected error in predicting the vibrational frequency by the ENMO-CI at its best is still on the order of a few per cent. Although such performance for vibrational frequencies is acceptable for some popular density functional methods [132], the contemporary electronic molecular orbital CI-type treatments allow one to obtain vibrational frequencies with the error of one wave number [133, 134, **?**] and should in principle converge to the exact result.

To corroborate the simple analysis shown above by direct computation and determine

the rate of the convergence to the estimated result with the number and type of particle excitations, we carried out a series of ENMO-CIS, ENMO-CIS[D], ENMO-CISD and ENMO-CISD[T] calculations of the low-lying vibrational states of the diatomics using the ENMO model with one of the nuclei in the molecule frozen. By CIS we mean that the CI wave function is constructed from all the possible single-particle excitations. CIS[D] includes all the single excited configurations as well as the double configurations composed of one-electron and one-proton excitations. CISD includes all the possible single and double excitations as well as CISD[T], but the latter also contains the triple excitations that are made of the promotions of two electrons and one proton only.

Calculations of the vibrational properties with the ENMO model have been done before. Thomas [135] calculated a few vibrational transitions of hydrogen fluoride, lithium hydride, water, methane and ammonia, but his method was quite *ad hoc* and the frequencies had only a qualitative agreement with experiment, typically with an error on the order of 10%. Nakai and co-workers [116] gave the vibrational spectra of the isotopomers of the hydrogen molecule by the ENMO-CIS method, but since their approach did not remove the spurious states from the spectra (for example, by fixing one of the nuclei in the molecule), we do not know how these authors recognized the ”true” states. Iordanov and Hammes-Schiffer [136] calculated numerically the ENMO Hessian in the classical degrees of freedom only and obtained the non-adiabatic corrections to the corresponding MO vibrational frequencies. This Hessian approach to vibrational analysis does not provide the vibrational frequencies that correspond to the movement of the quantum nuclei (thus, for the HCN molecule, in which both the carbon and the nitrogen atoms were treated classically, only the C-N vibrational frequency was obtained).

In what follows we show that the vibrational states that correspond to the movement of the quantum nuclei naturally arise from the ENMO-CI calculations, when one of the nuclei is frozen. The placement of the orbitals in our diatomics calculations is depicted in Figure 15. The electronic and nuclear s-orbitals can be put at several points along the

vibrational coordinate in order to model the delocalization of the vibrating nucleus.

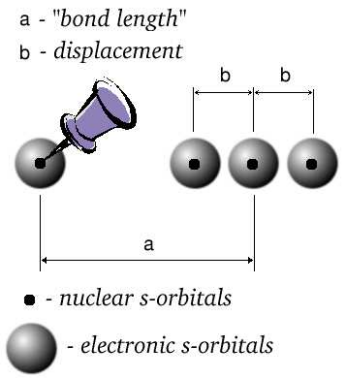


Figure 15: The placement of electronic and nuclear orbitals in diatomics for the description of the low-lying vibrational states. The thumb-tack indicates the fixed (classical) nucleus.

It is also possible to use the single-centered p-, d-, ... functions in the nuclear basis set but this will also describe librations in the directions perpendicular to the vibration coordinate and will lead to the degeneracies in the spectrum. Since our calculations have only a demonstrative character, we set the parameter a equal to the equilibrium bond length. However, it is not exactly clear how to choose the parameters b_i as well as how many of them to take. To obtain the quantitative result for the v-th vibrational state one may need to place the orbitals at more than $v + 1$ along the vibrational coordinate owing to the incompleteness of the basis sets used. We will be referring to placing the orbitals symmetrically at $2n + 1$ points in space as using n b-parameters.

The fewer nodes the wave function possesses the easier the description of the state it belongs to, therefore in our calculations we were chiefly striving to obtain the $v = 0$ and $v = 1$ energies. Because of the implicit inclusion of the anharmonicity effects in the ENMO scheme, it is not possible to obtain the harmonic vibrational frequency ω knowing only the energies of the states $v = 0$ and $v = 1$. The convenient characteristic that may be easily compared to the experimental result is the energy gap $G(10) = E_{v=1} - E_{v=0}$. As in the case of the vibrational frequencies, the gaps calculated from ω_∞ should be slightly higher than

those the ENMO-FCI model should produce in the limit of an infinite basis set because of the anharmonicities.

In calculating the energy gaps $G(ij)$ one can vary the type of the excitations included in the wave function (CIS, CIS[D], CISD, CISD[T]), the size of the electronic and nuclear basis sets, and the number and magnitude of the displacements b. Since we only wish to confirm numerically the prediction of the simple formula (93), it is not our intention here to find the best possible set of parameters for the calculation of $G(ij)$ and to present the exhaustive study of the way the vibrational spectrum behaves as these parameters change. Therefore we present only the most characteristic of our results in Tables 6-8 and summarize them with this discussion.

(i) We studied three diatomic hydrides: H_2, ^7LiH and $H^{19}F$.

(ii) The general character of results depends only slightly on the size of the basis set taken at the points along the vibrational coordinate. Therefore we present the numbers obtained at the moderate level [6-31G:4s], where the exponents for the 4s nuclear functions were taken from Ref. [115].

(iii) In all the cases studied the ground state has the largest contribution from the SCF reference determinant. The several next excited states are composed mainly of the nuclear determinants, which allows to ascribe to them the nuclear vibrational character.

(iv) As tables 6-8 indicate, the ENMO-CIS method generally overestimates $G(10)$ and gives a very poor agreement with either the experimental value or the value obtained from ω_∞ whatever the parameters are. ENMO-CIS[D] still overestimates $G(10)$ but not so much as ENMO-CIS. ENMO-CISD, seemingly surprisingly, instead of the further improvement of the result, makes it much worse. This phenomenon has a logical explanation, though. Let $\Phi_0 \equiv \Phi_0^e \Phi_0^n$ denote the SCF reference determinant. Then $\Phi_{ij...}^{eab...} \Phi_{pq...}^{nfg...}$ will represent the excited configuration formed from the excitations of electrons from the orbitals $i, j, ...$ to the orbitals $a, b, ...$ and nuclei from the orbitals $p, q, ...$ to the orbitals $f, g, ...$ It is clear that the single excitations only cannot describe correlation effects because the matrix elements

$\langle \Phi_i^{ea} | H | \Phi_0 \rangle$ and $\langle \Phi_p^{nf} | H | \Phi_0 \rangle$ are equal to zero. Therefore, in ENMO-CIS both the ground state and the first excited vibrational state do not include correlation. In ENMO-CIS[D], the matrix elements $\langle \Phi_i^{ea} \Phi_p^{nf} | H | \Phi_0 \rangle$ are generally not zero. The functions $\Phi_i^{ea} \Phi_p^{nf}$ are very likely to contribute to the vibrational excited states and, consequently, lower their energy which makes $G(10)$ decrease. Electronic-electronic correlation is expected to be much larger than the electronic-nuclear one, and that is why ENMO-CISD produces a very large gap $G(10)$ that is a difference between the energy of the ground state to which the functions Φ_{ij}^{eab} contribute greatly, and the energy of the first excited vibrational state, which does not have such contributions. ENMO-CISD[T] gives the best results since both the ground and the excited vibrational states include the doubly-excited electronic configurations that are responsible for the description of the most important part of the correlation. The values $G(10)$ obtained from ENMO-CISD[T] show the closest agreement with the prediction of formula (93).

(v) The values $G(21)$ and $G(32)$ roughly follow $G(10)$, but do not necessarily satisfy the relation $G(10) > G(21) > G(32)$. The correct description of the states $v = 2$ and $v = 3$ within the ENMO scheme apparently requires even larger basis sets and levels of particle excitation.

(vi) Among wave functions obtained with different values and number of parameters b, it is not clear which is 'the best one': while the basis set is incomplete the function that corresponds to the lowest energy does not have to provide the most adequate value of $G(10)$. Yet, having no other guide-line than energy, we must favor the gap that corresponds to the function with the lowest energy. Thus, the best values of $G(10)$ for H_2, LiH and HF that we were able to obtain are 2952 a.u. (four parameters b), 1402 a.u. (four parameters b), and 4622 a.u. (one parameter b), respectively. Perhaps, the best result for the HF molecule should be considered 3865 a.u. which is obtained with the ENMO-CIS[D] model, but with four parameters b.

(vii) Taking just one parameter b generally gives the greatly overestimated values of $G(10)$

Table 6: The ENMO-CI results for H_2 molecule. The parameters a and b are defined in figure 15 and are given in angstroms; $a = 0.7414$. The basis set is [6-31G:4s].

H_2								
Method	b_1	b_2	b_3	b_4	Ground state energy, a.u.	$G(10), \mathrm{cm}^{-1}$	$G(21), \mathrm{cm}^{-1}$	$G(32), \mathrm{cm}^{-1}$
ENMO-CIS	0.05	–	–	–	-1.08836941	5332	5112	2869
–”–	0.075	–	–	–	-1.08892457	5335	4872	3108
–”–	0.10	–	–	–	-1.08905442	6321	3999	3830
–”–	0.12	–	–	–	-1.08918459	7875	2555	4716
ENMO-CIS[D]	0.05	–	–	–	-1.09542972	3924	4885	2524
–”–	0.075	–	–	–	-1.09694130	3474	4978	2620
–”–	0.10	–	–	–	-1.09514467	3988	4835	2524
ENMO-CISD	0.05	–	–	–	-1.11930231	9150	4826	2527
–”–	0.075	–	–	–	-1.12069450	8671	4885	2684
–”–	0.10	–	–	–	-1.12075732	8559	5175	2199
ENMO-CISD[T]	0.05	–	–	–	-1.11993496	4148	4964	2577
–”–	0.075	–	–	–	-1.12166446	3707	5035	2695
–”–	0.10	–	–	–	-1.12207721	3698	5348	2221
–”–	0.06	0.02	0.02	0.02	-1.12222953	3237	3632	2956
–”–	0.08	0.02	0.02	0.02	-1.12228908	3088	3356	3595
–”–	0.10	0.02	0.02	0.02	-1.12304290	3031	3246	3533
–”–	0.12	0.02	0.02	0.02	-1.12316966	2985	3152	3373
–”–	0.14	0.02	0.02	0.02	-1.12346835	2952	3024	3160
–”–	0.16	0.02	0.02	0.02	-1.12333900	2983	3057	2976
Experiment	–	–	–	–	–	4157	3915	3672
From ω_∞	–	–	–	–	–	3112	3112	3112

for any level of excitation. At the ENMO-CISD[T] level of theory four parameters b produce the gaps $G(10)$ for H_2 and LiH that differ from the prediction of formula (93) by about 5%. Unfortunately, we were not able to take more than one displacement b for the HF molecule at the ENMO-CISD[T] level of theory because in this case the number of determinants exceeded the capabilities of our simple code.

5.8 Conclusions

As similar to the traditional MO approach as it may seem, the ENMO as a fully non-adiabatic model is built on substantially different principles that are quite new to the electronic structure theory. The most notable of them is the jettison of the potential energy surface - there is no longer such concepts in the ENMO approach as ‘equilibrium geometry’ and ‘reaction mechanism’. A single ENMO-CI computation should give the electronic, vibrational and (in principle) rotational spectrum of the molecule. The rotational spectrum,

Table 7: The ENMO-CI results for LiH molecule. The energetic values are presented in a.u. The parameters a and b are defined in figure 15 and are given in angstroms; $a = 1.5957$. The basis set is [6-31G:4s].

^{7}LiH								
Method	b_1	b_2	b_3	b_4	Ground state energy, a.u.	$G(10), \mathrm{cm}^{-1}$	$G(21), \mathrm{cm}^{-1}$	$G(32), \mathrm{cm}^{-1}$
ENMO-CIS	0.075	–	–	–	-7.94397307	4108	5412	3318
–”–	0.10	–	–	–	-7.94357498	3702	5956	2151
–”–	0.12	–	–	–	-7.94346108	5026	4675	1447
ENMO-CIS[D]	0.10	–	–	–	-7.95236298	2293	4389	2818
–”–	0.12	–	–	–	-7.95281874	2204	4567	2443
–”–	0.14	–	–	–	-7.95309296	2182	4435	2394
–”–	0.16	–	–	–	-7.95474381	2539	3979	2741
–”–	0.18	–	–	–	-7.95441832	2581	3393	3529
–”–	0.20	–	–	–	-7.95358820	2550	2965	4119
ENMO-CISD	0.075	–	–	–	-7.97179716	6599	4348	2987
–”–	0.1	–	–	–	-7.97253341	6705	4229	2976
–”–	0.12	–	–	–	-7.97194190	6202	4137	2787
ENMO-CISD[T]	0.10	–	–	–	-7.97358690	2590	4523	2785
–”–	0.14	–	–	–	-7.97426842	2504	4457	2487
–”–	0.18	–	–	–	-7.97385315	2555	3327	3817
–”–	0.10	0.02	0.02	0.02	-7.97502919	1906	2311	3817
–”–	0.14	0.02	0.02	0.02	-7.97586990	1659	1974	2278
–”–	0.18	0.02	0.02	0.02	-7.97774163	1532	1848	1877
–”–	0.22	0.02	0.02	0.02	-7.97752741	1386	1940	1416
–”–	0.18	0.04	0.04	0.04	-7.97790417	1402	1784	1601
–”–	0.20	0.04	0.04	0.04	-7.97777510	1335	1827	1395
–”–	0.22	0.04	0.04	0.04	-7.97729540	1254	1799	1352
Experiment	–	–	–	–	–	1393	1380	1367
From ω_∞	–	–	–	–	–	1314	1314	1314

Table 8: The ENMO-CI results for HF molecule. The parameters a and b are defined in figure 15 and are given in angstroms; $a = 0.91706$. The basis set is [6-31G:4s].

H^{19}F								
Method	b_1	b_2	b_3	b_4	Ground state energy, a.u.	$G(10), \mathrm{cm}^{-1}$	$G(21), \mathrm{cm}^{-1}$	$G(32), \mathrm{cm}^{-1}$
ENMO-CIS	0.075	–	–	–	-99.9533464	5386	4247	4673
–”–	0.10	–	–	–	-99.9536391	6330	3435	6556
–”–	0.125	–	–	–	-99.9540076	8092	1939	7982
ENMO-CIS[D]	0.075	–	–	–	-99.9598325	4249	4482	2550
–”–	0.10	–	–	–	-99.9602766	4365	4552	2827
–”–	0.125	–	–	–	-99.9603874	4673	4515	3483
–”–	0.06	0.02	0.02	–	-99.9601453	3898	4163	791.8
–”–	0.10	0.02	0.02	–	-99.9608956	3825	4065	1008
–”–	0.12	0.02	0.02	–	-99.9612192	3786	3937	1220
–”–	0.12	0.02	0.02	0.02	-99.9613549	3779	3900	1259
–”–	0.14	0.02	0.02	0.02	-99.9615122	3865	3977	1092
ENMO-CISD	0.075	–	–	–	-100.071872	28790	4464	2530
–”–	0.10	–	–	–	-100.071560	28710	4604	2748
–”–	0.125	–	–	–	-100.071061	28790	4673	3323
ENMO-CISD[T]	0.075	–	–	–	-100.073438	4486	4372	3132
–”–	0.10	–	–	–	-100.073868	4532	4541	3463
–”–	0.125	–	–	–	-100.073915	4622	4752	4025
Experiment	–	–	–	–	–	4132	4125	4118
From ω_∞	–	–	–	–	–	4032	4032	4032

however, would be extremely hard to obtain in practice, because the requirements for the basis set completeness are much more exacting for the description of the rotations.

By virtue of its technical similarity to the standard electronic structure theory methods, the ENMO approach is attractive for the non-BO description of molecular structure, and it has generated significant interest recently. However, a number of serious conceptual difficulties arise, and in this work we have pointed out some of these for the first time and have elucidated others. We show that since the ENMO ansatz (74) is not generally separable into 'internal' and the CM parts, even if one subtracts the center-of-mass kinetic energy from the Hamiltonian, the computed ENMO spectrum of the molecular Hamiltonian for the hydrogen atom is littered with spurious states. We conclude that the ENMO method in its present form is hardly useful for computing excited molecular states when all particles are treated quantum-mechanically.

It is thus clear that to circumvent the spurious states problem one has to either (i) formulate the problem fully in the internal coordinates, at the expense of technical simplicity [98] or (ii) break the translational invariance of the Hamiltonian. The latter might be achieved by putting a molecule into a potential well, similarly to the approach in nuclear structure computations. Here we chose a simpler route to avoid spurious states problem by fixing (i.e. describing classically) one heavy particle. However, we have discovered by simple analytical estimations and numerical ENMO-CI calculations that the hybrid ENMO-CI method, with one of the nuclei in the molecule fixed, gives errors of a few per cent in predicting low-lying vibrational states even for such 'favorable' molecules as HF. Furthermore, the selection of the basis set for nuclei described quantum-mechanically in ENMO methods remains ambiguous and a challenge for the general applicability of such methods to the computation of vibrational spectra.

Another conceptual problem with the ENMO-type ansatze is their inability to describe efficiently the interparticle cusps of the full wave function. This may translate in practice into slow asymptotic convergence of computed energies with respect to the one-particle

basis set size. Thus, for the hydrogen atom, we observe that the basis set incompleteness error in the largest ENMO calculation we report here is an order of magnitude greater than in the corresponding BO computation.

The aforementioned fundamental issues, and our experience with the ENMO methods discussed here cast doubt on the possibility that they may serve as highly accurate, truly non-adiabatic models models of molecular systems when all nuclei are treated quantum-mechanically. However, by describing only light nuclei quantum-mechanically, the spurious states problem may be circumvented. Thus it may be possible to use the ENMO method for chemically relevant non-adiabatic, tunnelling, isotopic, etc. problems. The important yet still unanswered question is how the inefficient description of the interparticle cusps by the ENMO wave functions will affect the practical performance of ENMO methods in such applications. It is clear that more testing and empirical evidence are needed before we can judge the usefulness of ENMO methods. We predict that the most important future advances in ENMO models will be probably those which solve the problems raised by the present study.

The material of this chapter was published in *Molecular Physics* [137].

CHAPTER VI

A GENERAL DIAGRAMMATIC ALGORITHM FOR CONTRACTION AND SUBSEQUENT SIMPLIFICATION OF SECOND-QUANTIZED EXPRESSIONS

6.1 Introduction

The rapid development and active usage of *ab initio* methods that include electron correlation have significantly changed the face of quantum chemistry that heretofore had been focusing mainly on the Hartree-Fock self consistent field (SCF) theory and its semi-empirical extensions. The MO theory provides the one-electron basis for the vast majority of the post-SCF approaches, most popular of which are various flavors of the many-body perturbation (MBPT), configuration interaction (CI) and coupled cluster (CC) theories. The correlation in such methods is introduced through excitations from a relatively small number of occupied to a large pool of virtual orbitals, and as the number of various matrix elements which must be evaluated in these approaches rises sharply with the maximum excitation level, their efficient implementation calls for a number of ingenious numerical and computational simplification techniques. Apart from the colossal number of computational operations needed to produce final numbers, there is another serious hindrance that currently bars researchers from performing high-accuracy quantum chemical calculations routinely. These days, the equations of quantum chemical models may contain a very large number of algebraic terms, so that expressions are difficult, if at all possible, to derive, analyze and program. The idea to automate the generation and analysis of the

mathematical expressions as well as produce computer code took real shape in the work of Janssen and Schaefer [138] (although it was by no means the very first among the works in which analytical calculations were partly performed by computer programs – see, for example Ref. [62]). It was followed by a number of programs which were, however, often aimed at some specific class of quantum chemical model, usually coupled cluster theory [139, 140, 141, 142, 143, 144, 145, 146, 147, 148]. Recently a quite general program, the Tensor Contraction Engine (TCE) [149] that became a part of a much larger and ambitious project was developed by a number of researchers. A recent review by Hirata [150] describes the program's capabilities in detail. TCE, capable of handling in principle any algebraic theory, exploits Wick's theorem in order to derive the corresponding equations and then generates efficient computer code for the multiplication of the multidimensional arrays. However, TCE evaluates matrix elements that have a fixed form, which, although flexible enough to cover most of the currently used electronic structure theories, does not allow one to work with the matrix elements of arbitrary operators. Besides, TCE does not handle second-quantization operators explicitly, which is desirable in order to facilitate the construction of quantum chemical models which are not always written in terms of matrix elements.

Here we present a completely general algorithm to manipulate second-quantized expressions, whether scalars or operators. Our approach is based solely on the Goldstone diagrammatic technique [38], popular among method developers. Diagrams, of which those of the Goldstone type are the most elementary and transparent, serve as a visual and topological (rather than algebraic) representation of second-quantized expressions, and proved to be extremely useful in many areas of quantum chemistry. Our algorithm can also be extended to treat Hugenholtz-type diagrams [151] which are essentially of the same type as Goldstone diagrams, but whose vertices absorb antisymmetrization. Hugenholtz diagrams might therefore be used to produce compact algebraic expressions that feature permutation operators, but Goldstone diagrams are advantageous when programmable expressions are

needed or when the spin integration of the obtained expressions is planned. Also deserving of consideration are the spin-free diagrams of Kutzelnigg [152] which combine the assets of both Hugenholtz and Goldstone diagrams.

Several excellent introductory sources exist that describe in great detail the Goldstone diagrammatic technique [36, 33, 34, 35, 153] and it would seem excessive to outline it in this work. However, so many various types of diagrams emerged since their first usage in quantum chemistry (some of which look almost identical), that, for the benefit of the reader, we wish to briefly describe the notation we employ in drawing diagrams throughout this chapter.

Each Goldstone diagram is a pictorial representation of the second-quantized expression for a given Fermi-vacuum, so that each general index that belongs to the spin-orbital creation or annihilation operators is either a particle (p) or hole (h) state index. Particle (hole) states lie above (below) the Fermi-vacuum. In this notation the particle (hole) creation operator is $a_p^\dagger$ (a_h), and the particle (hole) annihilation operator is a_p ($a_h^\dagger$). The diagrams we are using do not absorb numerical factors (which emanate from the definition of the second-quantized operators or from addition of one or more identical diagrams) to avoid confusion. They also do not include antisymmetrization, so that the expression for the closed shell SCF energy

$$E_{SCF} = \sum_h h_{hh} + \frac{1}{2}\sum_{hh'}(g_{hh',hh'} - g_{hh',h'h}) \tag{94}$$

will be represented diagrammatically as shown in Figure 16. The first diagram (a 'bubble') in it represents the core energy, the second (a 'double bubble') – the electron-electron Coulomb energy, and the third (an 'oyster') – the electron exchange energy.

The central idea behind the development of our program was to make it possible for the researcher (user) to perform any permissible operations on diagrams automatically. In other words, we would like to create a development tool that would be an aid in the manipulation and analysis of second-quantized expressions through which many quantum-chemical

models can be written. In this work we present only the algorithm of our diagrammatic approach and do not give many details about its specific implementation, since the algorithm may be realized in various computer languages. Our current version of *Nostromo* is written in the symbolic package Mathematica [154].

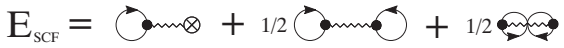


Figure 16: The expression for the SCF energy in Goldstone diagrammatic notation.

Nostromo works only with time-independent, discrete-basis and particle-conserving expressions. Diagrammatically, after the Fermi vacuum and the basis have been chosen, any such expression will be represented by one or more vertices and lines, coming in and out of these vertices. There will be vertices of four types only: particle creation and particle annihilation, particle creation and hole creation, particle annihilation and hole annihilation, and, finally, hole creation and hole annihilation. There will also be four types of lines – those that face down and are located below a vertex (so called down-below lines), and, using the analogous nomenclature, down-above, up-below and up-above lines (see Figure 17).

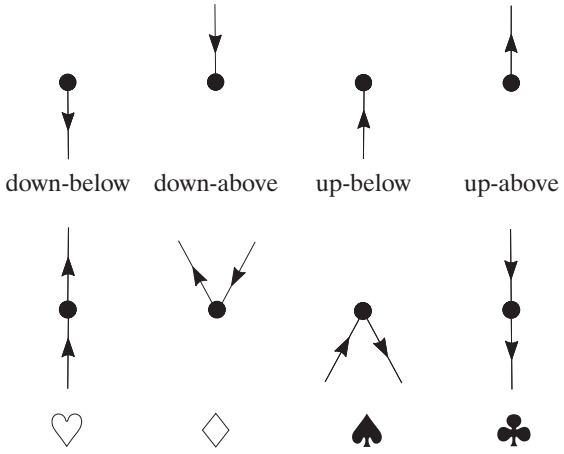


Figure 17: Types of lines and vertices in Goldstone diagrams.

6.2 Textual representation of Goldstone diagrams

The unifying idea of *Nostromo* is that each diagram may be 'encoded' by a text string, and vice versa, each such string of code may be translated back into a pictorial representation of the second-quantized expression – a Goldstone diagram. We call such a string of text a *textual diagram*. Every essential part of the diagram is represented by certain symbols in the text. In principle, this coding system does not provide a one-to-one correspondence between the pictures and algebraic formulae, but this is not its fundamental flaw and merely reflects the fact that there exists no one-to-one correspondence between algebraic expressions and diagrams. For example, an algebraic expression may be written as a diagram or as this diagram's mirror reflection, which do not always coincide. Conversely, a diagram that represents the multiplication of two vectors is translated back into the algebraic language equivocally – either as $\sum_i A_i B_i$ or as $\sum_i B_i A_i$. However, our coding system guarantees the isomorphic correspondence, so that any operation performed on the algebraic expression may be unequivocally conducted on its textual representation.

It should be stressed from the very start that diagrams (and, consequently, textual diagrams) are written only for normally-ordered second-quantization expressions. Therefore, the analog of n-particle operator in second quantization may be represented by Goldstone diagrams that have not only $2n$ open lines but also those that have $2n-2, 2n-4, ..., 0$ open lines. The reduction in the number of lines results from the contractions which appear as a 'penalty' for writing the operator in normal form. The textual diagram contains the names of the operators that constitute this diagram, as well as the information about how the different operators are connected by the particle and hole lines. Each text string starts with the name of an operator. It may be any operator which is included in this diagram. The name is always made of characters enclosed within two open square symbols (□). For example, □g□ means the name 'g'. The four types of vertices mentioned in the introduction are denoted by the following four symbols, respectively: ♡, ♢, ♠, ♣ (see Figure 17). After each such suit goes the information about the two lines that belong to the corresponding vertex.

The line that leaves the vertex always goes first. It is necessary to show whether the line bears any index or if it is summed over. The line may be connected with another line, and it should be seen from the information in the text with which one. The information about each line is put between two delineators – an opening $\pounds$ and a closing \$. The line that faces up and which does not have any index (that is, summed over) is denoted by the symbol $\uparrow$. Similarly, the line that faces down and is summed over, is written as $\downarrow$. For example, the one-particle operator that includes all the particle creation and annihilation operators

$$\hat{h} = \sum_{pp'} h_{pp'} a_p^\dagger a_{p'} \tag{95}$$

is written in *Nostromo* in the following simple way: $\square h \square \heartsuit \pounds \uparrow \$ \pounds \uparrow \$$. The one-particle operator in which the summation runs over all the indices – particle and hole,

$$\hat{h} = \sum_{pp'} h_{pp'} \{a_p^\dagger a_{p'}\} + \sum_{ph} h_{ph} \{a_p^\dagger a_h\} + \sum_{hp} h_{hp} \{a_h^\dagger a_p\} + \sum_{hh'} h_{hh'} \{a_h^\dagger a_{h'}\} + \sum_{hh} h_{hh} \tag{96}$$

with the curly brackets indicating the normal ordering, is written in *Nostromo* as the sum:

$$\begin{aligned} &\square h \square \heartsuit \pounds \uparrow \$ \pounds \uparrow \$ \;\; + \;\; \square h \square \diamondsuit \pounds \uparrow \$ \pounds \downarrow \$ \;\; + \\ &\square h \square \spadesuit \pounds \downarrow \$ \pounds \uparrow \$ \;\; + \;\; \square h \square \clubsuit \pounds \downarrow \$ \pounds \downarrow \$ \;\; + \\ &\square h \square \clubsuit \pounds \backsim 1 \$ \pounds \backsim 1 \$. \end{aligned} \tag{97}$$

The last term has a contraction and its structure in our coding system is described in the next paragraph. Further, suppose that we want to encode the operator

$$\hat{h} = h_{pp'} a_p^\dagger a_{p'} \tag{98}$$

which has no summation over the indices. Then, in *Nostromo* coding system, we simply place the indices after the $\uparrow$ and $\downarrow$ symbols. Therefore, the above-mentioned operator will be written as $\square h \square \heartsuit \pounds \uparrow p \$ \pounds \uparrow p' \$$.

The contraction of the line is denoted by the symbol $\backsim$ followed by the contraction number, which replaces the arrow. For example, the contracted operator

$$\hat{h}_{Sc} = \left(\sum_{ij} h_{ij} a_i^\dagger a_j \right)_{Sc} = \sum_h h_{hh}, \tag{99}$$

where the subscript Sc denotes the scalar expression, is encoded as $\square h \square \clubsuit \pounds \backsim 1\$ \pounds \backsim 1\$$. If the contraction retains the indices, as in

$$\left(h_{hh} a_h^\dagger a_h \right)_{Sc} = h_{hh}, \tag{100}$$

its code becomes $\square h \square \clubsuit \pounds \backsim 1h\$ \pounds \backsim 1h\$$.

In case all the matrix elements $h_{ij\ldots} = const$, that is, the operator does not have a name, 1 is inserted between two open squares. For example,

$$\sum_{pp'} a_p^\dagger a_{p'} \tag{101}$$

is coded by $\square 1 \square \heartsuit \pounds \uparrow \$ \pounds \uparrow \$$.

Using the described *Nostromo* coding system, it is possible to encode any complicated second-quantized expression. The linear combination of second quantized expressions is written as a linear combination of textual diagrams.

6.3 Nostromo functions

The delineated encoding system bridges the gap between the pictorial representation of a second-quantized expression, which is appealing to a human, and the minimum amount of information needed for this second-quantized expression to be unequivocally recognized by a computer interpreter. It is straightforward now to write functions that work with diagrams by analyzing and parsing the textual diagrams.

Nostromo includes primitive functions, such as those that count the number of specific lines in a textual diagram; auxiliary functions that create textual diagrams (DiagramGenerator) or select them according to some criteria (for example, by the number of uncontracted

lines, connectedness etc.); and complicated, primary functions such as those that perform all the possible contractions among multiple diagrams (MultiContractor) and translate the received output (Translator) into the simplest algebraic expression, taking into account permutational and other types of symmetries. The automatic generation of diagrams occupied the attention of researchers in the past [155, 156, 157], but these works were mostly concerned with fully-contracted diagrams which appeared in perturbation expansions. DiagramGenerator produces all the topologically non-equivalent diagrams that constitute an n-particle operator, and in combination with the selection functions and MultiContractor, is capable of generating, in principle, any types of diagrams. While DiagramGenerator is easy to devise and implement, we nevertheless intend to characterize it in some detail in the next section, in order to render our chapter more complete and our discussion of the subsequent material clearer. In comparison with DiagramGenerator, the functions MultiContractor and Translator have a rather complicated structure. MultiContractor is solely topological and does not directly employ Wick's theorem. The Translator function solves the important problem of identifying equivalent terms in the algebraic expression and their subsequent factorization. Instead of using a canonical form of any kind, it analyzes the topological structure of each diagram and equates terms with the same topological structure. The detailed description of these two functions will be given in sections 5 and 6.

6.4 The DiagramGenerator function

We begin with writing a general n-particle operator $Q^{(n)}$ (in which all n particles are equivalent) in second quantization form without explicitly specifying the Fermi level:

$$Q^{(n)} = \frac{1}{n!} \sum_{i_1 i_2 ... i_n} \sum_{j_1 j_2 ... j_n} Q_{i_1 i_2 ... i_n, j_1 j_2 ... j_n} \times a^\dagger_{i_1} a^\dagger_{i_2} ... a^\dagger_{i_n} a_{j_n} a_{j_{n-1}} ... a_1. \qquad (102)$$

As we introduce the Fermi level, the indices $i_1 i_2 ... i_n$ take on additional labels p (particle state) and h (hole state), and the operator $Q^{(n)}$ may be written as the sum of 2^{2n} terms, each of which is a summation analogous to (102) with the indices carrying distinct labels

p or h. If we wish to work with $Q^{(n)}$ in a diagrammatic representation, we apply the Wick's theorem to it, so that it becomes the sum of the above-specified $2^{2n} = 4^n$ terms, each brought to the normal order, plus a number of additional sums (also written in normal order) which contain one or more contractions – so called 'bubbles' and 'oysters'. It is easy to see that each of the uncontracted sums may be written diagrammatically as a series of n vertices positioned along a horizontal line. Each vertex can be any of ♡, ♢, ♠, ♣, totaling exactly 4^n different diagrams. As the vertices of a given diagram may be put in any order because of the particles' equivalence, the number of distinct diagrams may be reduced by bringing each of the 4^n uncontracted diagrams to some canonical form, for example, that in which all the vertices are arranged in the following order: ♡, ♢, ♠, ♣. A similar canonicalization is possible for the contracted terms. Concurrently to the described process, the DiagramGenerator function generates 4^n possible combinations of the four vertices for a n-particle operator, and produces a sum of the corresponding textual diagrams. Then, it adds to this sum all the textual diagrams with $1, 2, ..., n$ contractions, thus forming a representation of the n-particle operator in normal form. After that, the vertices are rearranged according to the chosen canonical order. The process of generating diagrams, rearranging their vertices and reducing the number of terms is illustrated for a two-particle operator $\hat{g}$ in Figure 18.

6.5 The MultiContractor function

As an input, MultiContractor receives a list of textual diagrams $\{A_1, A_2, ..., A_M\}$ as well as optional parameters that impose restrictions on the number or type of contractions to be performed. To give an example of the latter, we may mention a very often desired restriction that only fully contracted diagrams should be generated. An ordered sequence or list of (textual) diagrams $\{A_1, A_2, ..., A_M\}$ will be called an (uncontracted) configuration, with A_1 corresponding to the operator $\hat{A}_1$, A_2 corresponding to the operator $\hat{A}_2$, and so on. The order of diagrams in $\{A_1, A_2, ..., A_M\}$ is the same as the order of operators $\hat{A}_1, \hat{A}_2, .., \hat{A}_M$

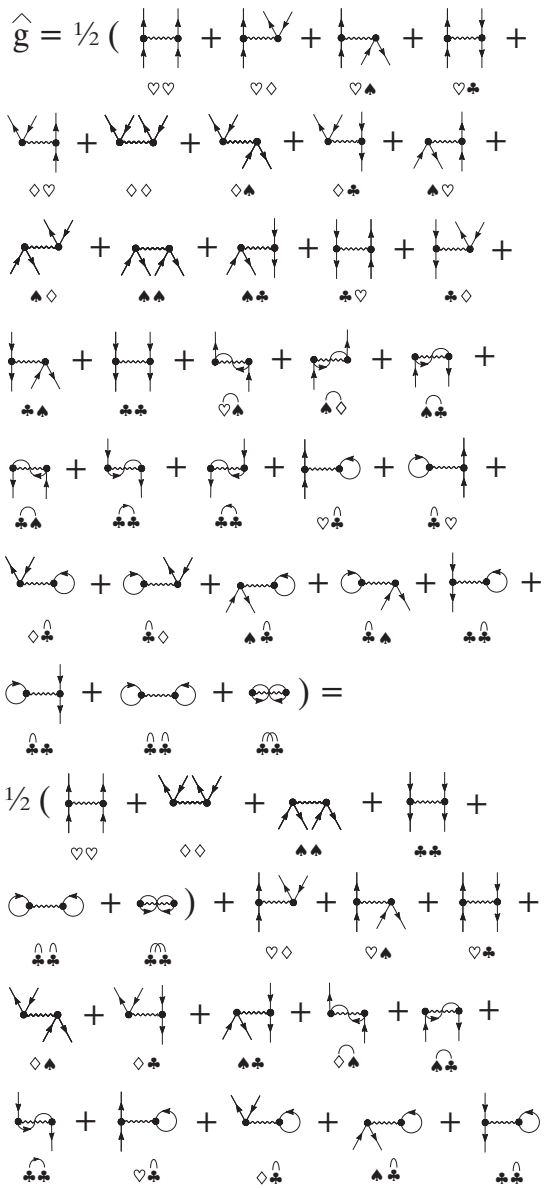


Figure 18: The general two-particle operator in Goldstone diagrammatic notation: reducing the number of diagrams by identifying symmetry-equivalent ones.

in the product $\hat{A}_1\hat{A}_2...\hat{A}_M$. If one or more contractions are performed among the diagrams of the sequence $\{A_1, A_2, ..., A_M\}$, the configuration is called a contracted configuration. At the outset, in order to make the explanations more clear, we shall assume that none of the

above-mentioned restrictions are imposed and that all the possible contractions, conforming to a set of rules R, are allowed. MultiContractor realizes the following contraction rules R: ($R1$) The diagrams to be contracted are positioned from above to below with the uppermost diagram corresponding to the leftmost textual diagram A_1 in the list and the lowermost diagram corresponding to the rightmost textual diagram A_M. ($R2$) Up-lines are connected with up-lines and down-lines are connected with down-lines only. ($R3$) Lines cannot be connected more than one time. ($R4$) Any up (down)-below line that belongs to the diagram A_k may be connected with any up (down)-above line that belongs to the diagram A_l only if $k < l$. Any other connections are prohibited. Note that up-lines are contracted independently from down-lines and that the set of all the possible contractions for a given configuration becomes a direct product of the set of up-contractions and down-contractions.

Suppose a configuration of M textual diagrams $\{A_1, A_2, ..., A_M\}$ is passed to the MultiContractor function together with the demand to generate, as a result of (possible) contractions, only those textual diagrams that have a specified number of up-below, up-above, down-below, and down-above lines. The number of performed contractions will be denoted as C. The following describes the MultiContractor algorithm.

(i) Calculate the number of up-above, up-below, down-above and down-below lines (denoted as N_k^{ua}, N_k^{ub}, N_k^{da}, and N_k^{db} respectively) for each diagram in $\{A_1, A_2, ..., A_k, ..., A_M\}$. Also calculate the total number of these lines. For example, the total number of up-above lines in the given configuration is $N^{ua} = \sum_{k=1}^{M} N_k^{ua}$.

(ii) In case MultiContractor is to perform the average value calculation, it need not generate diagrams with unconnected lines since they give zero contribution to the final result. The necessary condition for the list of diagrams to produce at least one fully-contracted diagram is furnished by these equalities

$$N_1^{ua} = N_1^{da} = 0, \qquad N_M^{ub} = N_M^{db} = 0, \tag{103}$$

$$\sum_{i=1}^{M-1} N_i^{ub} = \sum_{i=2}^{M} N_i^{ua}, \qquad \sum_{i=1}^{M-1} N_i^{db} = \sum_{i=2}^{M} N_i^{da}. \tag{104}$$

Requirements (103) are a direct consequence of the rule ($R4$). If any of the statements (103, 104) is not satisfied, MultiContractor instantly evaluates the contribution to the average value to be zero. Such a 'forecasting' allows to accelerate the calculation of scalar expressions significantly.

(iii) Calculate the maximal number of up-above, up-below, down-above and down-below lines which can possibly be contracted (denoted as C_{max}^{ua}, C_{max}^{ub}, C_{max}^{da}, and C_{max}^{db} respectively). Obviously, $C_{max}^{ua} = C_{max}^{ub}$ and $C_{max}^{da} = C_{max}^{db}$, therefore we shall only operate with $C_{max}^{u} \equiv C_{max}^{ua}$ and $C_{max}^{d} \equiv C_{max}^{da}$. Then $C^u = 0, 1, 2, ..., C_{max}^u$ and $C^d = 0, 1, 2, ..., C_{max}^d$.

Now we shall show how to calculate C_{max}^u. The same arguments are applicable without any change to the calculation of C_{max}^d. For a k-th diagram, $1 \leq k \leq M-1$, calculate $\gamma_k^u = \sum_{i=2}^{k} N_i^{ua} - \sum_{i=1}^{k-1} N_i^{ub}$. If $\gamma_k^u \geq 0$ (which means that the below-lines belonging to the diagrams with number l, $1 \leq l \leq k-1$, may be fully contracted with the above-lines belonging to diagrams with number m, $2 \leq m \leq k$), the number of the above-lines with which the below-lines of the diagram k may be contracted, is $Q_k^u = \min[\sum_{i=k+1}^{M} N_i^{ua}, N_k^{ub}]$. If $\gamma_k^u < 0$, $Q_k^u = \min[\gamma_k^u + \sum_{i=k+1}^{M} N_i^{ua}, N_k^{ub}]$. Thus, Q_k^u is the maximal number of up-contractions for the k-th diagram, given that all the below-lines belonging to the diagrams with numbers l, $1 \leq l \leq k-1$, are maximally contracted. C_{max}^u is obtained by summing over all Q_k^u's:

$$C_{max}^u = \sum_{k=1}^{M} Q_k^u. \tag{105}$$

In a similar manner,

$$C_{max}^d = \sum_{k=1}^{M} Q_k^d. \tag{106}$$

(iv) Now we move on to building the *map* $\mathcal{M}$ of all the possible contractions, which is simply a data array containing the information of how the lines can be connected. Since

up- and down-lines are connected independently, the map $\mathcal{M}$ is obtained by direct product of the up-map $\mathcal{M}^u$ and the down-map $\mathcal{M}^d$:

$$\mathcal{M} = \mathcal{M}^u \times \mathcal{M}^d. \tag{107}$$

The maps $\mathcal{M}^u = \{\mathcal{M}^u_j\}$ and $\mathcal{M}^d = \{\mathcal{M}^u_k\}$ are built in precisely the same way, therefore we shall only expound on how one builds $\mathcal{M}^u$.

Suppose that each up-below line of the configuration $\{A_1, A_2, ..., A_M\}$ is given an index from 1 to N^{ub} (starting from the leftmost line of the uppermost diagram and ending at the rightmost line of the lowermost diagram). Then an array of N^{ub} elements, called an up-below *billet* and denoted B^{ub} will indicate the contraction state of the up-below lines. The analogous numbering may be performed on the up-above lines, and an up-above billet B^{ua} containing N^{ua} elements may also be constructed. If none of the up-below lines is contracted with the up-above lines, the billets B^{ub} and B^{ua} are filled with zeros. Should the k-th up-below line be contracted with the l-th up-above line, B^{ub} and B^{ua} will become

$$B^{ub} = \{0, 0, ..., 1, 0, 0, 0, ...\}, \tag{108}$$

$$B^{ua} = \{0, 0, 0, ..., 1, 0, 0, ...\} \tag{109}$$

where 1 appears in B^{ub} in the k-th position and in B^{ua} in the l-th position. Similarly, should n up-contractions be performed, each of the billets B^{ub} and B^{ua} will contain n non zero-elements: $1, 2, ..., n$ at the positions that indicate the positions of the contracted lines in the diagrams. The integers $1, 2, ..., n$ serve as the dummy indices to show the structure of the contraction, and the combination of the billets B^{ub} and B^{ua} may represent an up-contraction. The list of all the legitimate up-contractions $\{[B^{ub}, B^{ua}]_i\}$ may be brought to the form $\{B^{ub}_i, \{B^{ua}_{ij}\}\}$ where every possible up-below billet B^{ub}_i is associated with the list of all the up-above billets, every one of which, if combined with B^{ub}_i, represents a valid up-contraction. The composite array that consists of elements $\mathcal{M}^u_i = [B^{ub}_i, \{B^{ua}_{ij}\}]$ is the up-map $\mathcal{M}^u$; (v) Since the up-map $\mathcal{M}^u$ includes all the up-contractions, its elements may

be grouped into sub-arrays according to the number of connections between below- and above lines. Thus, we shall have the sub-array representing zero contractions ${}^{(0)}\mathcal{M}^u$ that consists of one element, as well as other sub-arrays ${}^{(1)}\mathcal{M}^u$, ${}^{(2)}\mathcal{M}^u$, ..., ${}^{(C^u_{max})}\mathcal{M}^u$.

Now we turn our attention to constructing the legitimate contractions $\{B^{ub}_i, \{B^{ua}_{ij}\}\}$. Initially, we select the legitimate below-billets $\{B^{ub}_i\}$, and then, for each i, we select such above-billets $\{B^{ua}_{ij}\}$ that would form legitimate contractions, if combined with B^{ub}_i. For implementation purposes, it is convenient to build ${}^{(0)}\mathcal{M}^u$ first, then move on to ${}^{(1)}\mathcal{M}^u$ etc., and in the end, when all these sub-maps are generated, combine them into $\mathcal{M}^u$. There are a number of conditions (we call them below-conditions) that a below-billet must satisfy to be legitimate: (a) It must have N^{ub} elements – zeroes and distinct integers from 1 to C^u; (b) Below-lines that belong to the lowermost diagram should never be occupied, therefore the billet should have zeroes in the places that correspond to these lines; (c) This condition puts restrictions on the number of contractions the below-lines of each operator may have. As all the particles are equivalent, all the below lines are treated on equal footing, and if any restrictions apply as to whether the lines of the given operator may be contracted, these restrictions should concern the total number of contractions and not the specific lines. For each diagram number k, the number of its below-lines which may be possibly contracted in general depends on whether the below-lines of the diagrams with numbers l, $1 \leq l < k$ are already contracted. It may happen, for example, that some or all of the above-lines of the operators with numbers m, $k+1 \leq m \leq M$ are already contracted with the below-lines of the diagrams with numbers l, $1 \leq l < k$. The problem at this stage is that given only the above-billet we cannot know for sure *with which* above-lines the below-lines are contracted, we know only *which* below-lines are contracted. Therefore we select those below-billets which correspond to at least one legitimate contraction. For each of the diagrams (except the lowermost one) we check whether its below lines can possibly have as many connections as claimed for it in the given below-billet. The maximal number of the above-lines with which the below-lines of each diagram k may be connected

(this number cannot exceed the number of connections indicated in the below-billet for the k-th diagram if the billet is to be legitimate) is defined by the following scheme. Calculate $\lambda_k^u = \sum_{l=t+1}^{k} N_l^{ua} - \sum_{l=1}^{k-1} N_l^{ub}(\mathrm{conn})$, where t is the first diagram whose below-line is connected (the corresponding element in the billet is not zero) and $N_l^{ub}(\mathrm{conn})$ is the number of connected below-lines that belong to diagram number l (this information is extracted from the billet). If $\lambda_k^u \geq 0$ (which, similarly to the case of γ_k^u, means that the below-lines belonging to the diagrams with number l, $1 \leq l \leq k-1$, are fully contracted with the above-lines belonging to diagrams with number m, $2 \leq m \leq k$), then the maximal number of contractions for the below-lines of diagram number k is $P_k^u = \min[\sum_{i=k+1}^{M} N_i^{ua}, N_k^{ub}]$, and, consequently, the number of uncontracted below-lines for diagram k cannot exceed $N_k^{ub} - P_k^u$. This is the sought-for condition. However, if $\lambda_k^u < 0$, P_k^u is calculated according to the updated formula: $P_k^u = \min[\lambda_k^u + \sum_{i=k+1}^{M} N_i^{ua}, N_k^{ub}]$; (d) The billet should also comply with the optional restrictions on the number or type of contractions to be performed, which are given as input parameters at every MultiContractor function call; (e) Finally, to avoid multiple counting of contractions, the non-zero integers in the below-billets are placed in the strictly ascending order. All these conditions must be satisfied before the below-billet is accepted as valid. In an actual calculation, we generate all the billets that satisfy the criteria (a) and (e), and then select those that meet the criteria (b) and (c) and (d).

Once the below-billets $\{B_i^{ub}\}$, have been selected, it is easy to generate the above-billets $\{B_{ij}^{ua}\}$ for every i. The conditions for a legitimate above-billet are the following: (a) It must have N^{ua} elements – zeroes and distinct integers from 1 to C^u; (b) If contraction number k in the below-billet B_i^{ub} belongs to diagram number l then the contraction number l in all the above-billets $\{B_{ij}^{ua}\}$ should belong to the diagram number m, where $m > l$. This condition guarantees that no below-line is connected to an above-lying above-line. The fact that no restriction is laid down on the order of non-zero integers reflects the existence of permutations in many-body quantum theories. Examples of maps generated

by MultiContractor are shown in Table 9 and illustrated by Figure 19.

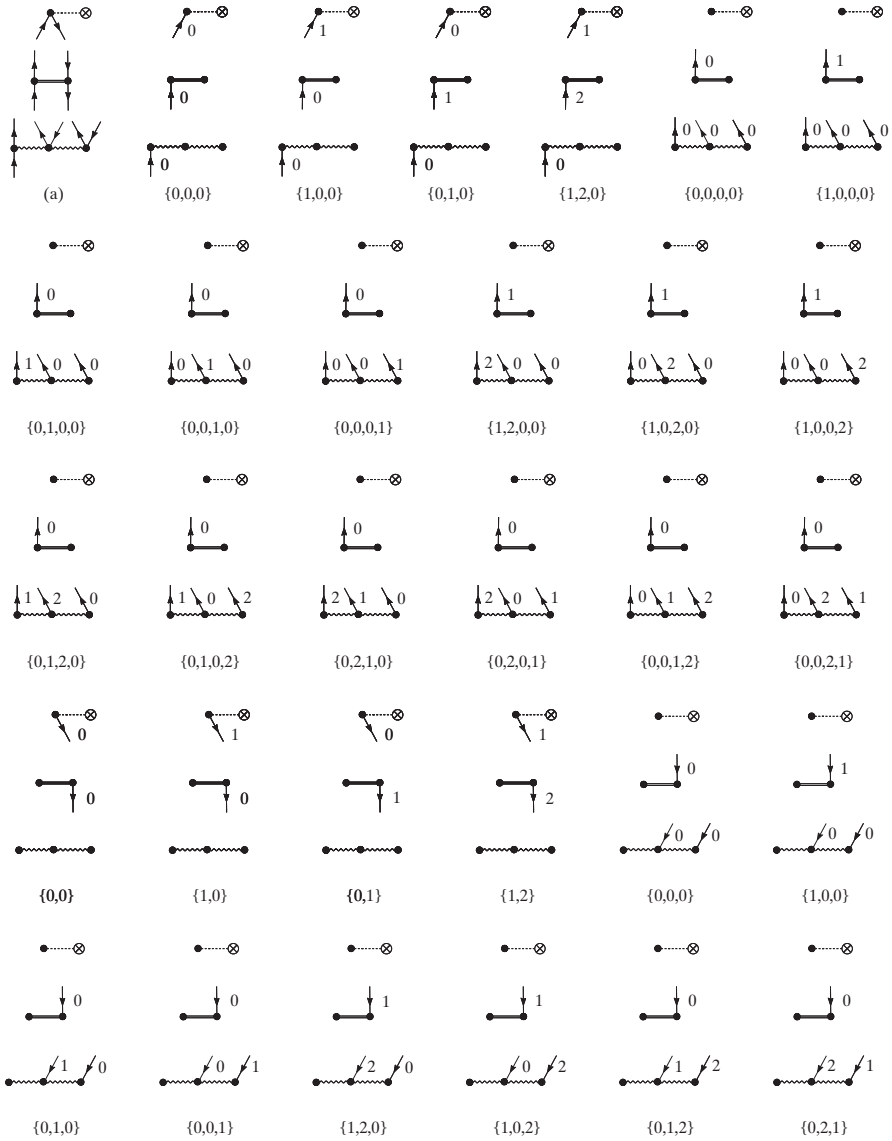


Figure 19: The billets of the configuration (a) and the corresponding numbering of lines.

.

(vi) Given the configuration $\{A_1, A_2, ..., A_M\}$ and its map $\mathcal{M}$, it is now possible to generate all the contracted textual diagrams by converting the open line symbols $\uparrow$ and $\downarrow$, at the locations specified by $\mathcal{M}$, into the contraction symbols $\backsim k$, where k is the contraction number. In case any of the diagrams in $\{A_1, A_2, ..., A_M\}$ come already partially or fully contracted with the maximal contraction number $k_{\max}$, the internal contractions numbers

Table 9: The map $\mathcal{M}$ generated by MultiContractor when applied to the configuration $\{\Box h \Box \spadesuit \pounds \downarrow \$\pounds \uparrow \$, \Box f \Box \heartsuit \pounds \uparrow \$\pounds \uparrow \clubsuit \pounds \downarrow \$\pounds \downarrow \$, \Box g \Box \heartsuit \pounds \uparrow \$\pounds \uparrow \$ \diamondsuit \pounds \uparrow \$\pounds \downarrow \$ \diamondsuit \pounds \uparrow \$\pounds \downarrow \$\}$ with no restrictions on the number or type of contractions. The configuration and the billets are illustrated by Figure 19. k is the number of contractions.

	${}^{(k)}\mathcal{M}^u$		${}^{(k)}\mathcal{M}^d$	
k	Below-billets	Above-billets	Below-billets	Above-billets
0	$\{0,0,0\}$	$\{0,0,0,0\}$	$\{0,0\}$	$\{0,0,0\}$
1	$\{1,0,0\}$	$\{\{1,0,0,0\},\{0,1,0,0\},\{0,0,1,0\},\{0,0,0,1\}\}$	$\{1,0\}$	$\{\{1,0,0\},\{0,1,0\},\{0,0,1\}\}$
	$\{0,1,0\}$	$\{\{0,1,0,0\},\{0,0,1,0\},\{0,0,0,1\}\}$	$\{0,1\}$	$\{\{0,1,0\},\{0,0,1\}\}$
2	$\{1,2,0\}$	$\{\{1,2,0,0\},\{1,0,2,0\},\{1,0,0,2\},\{0,1,2,0\},\{0,1,0,2\},\{0,2,1,0\},\{0,2,0,1\},\{0,0,1,2\},\{0,0,2,1\}\}$	$\{1,2\}$	$\{\{1,2,0\},\{1,0,2\},\{0,1,2\},\{0,2,1\}\}$

should be shifted by the $k_{\max}$, so that the dummy indices do not overlap. The textual diagrams with the substituted symbols $\uparrow$ and $\downarrow$ are concatenated to produce the final result. An example of the MultiContractor final result is given in Table 10 and illustrated by Figure 20.

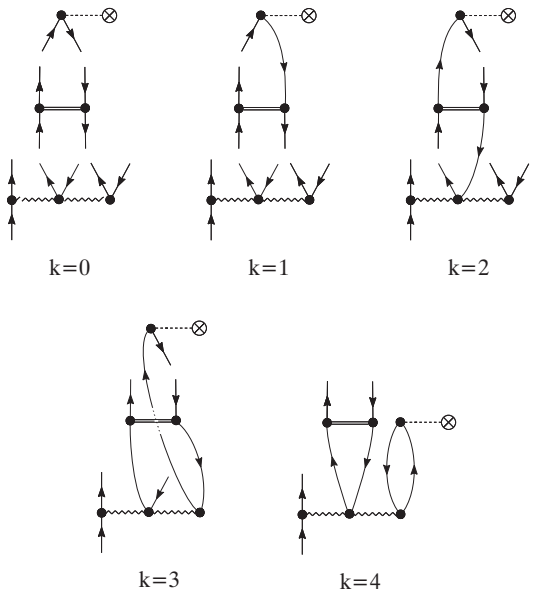


Figure 20: The diagrams matching the textual diagrams from Table II, where k is the number of contractions.

.

Table 10: Some of the textual diagrams produced as a result of applying MultiContractor to the configuration $\{\square h \square \spadesuit \pounds \downarrow \$\pounds \uparrow \$,\ \square f \square \heartsuit \pounds \uparrow \$\pounds \uparrow \clubsuit \pounds \downarrow \$\pounds \downarrow \$,\ \square g \square \heartsuit \pounds \uparrow \$\pounds \uparrow \$\diamondsuit \pounds \uparrow \$\pounds \downarrow \$\diamondsuit \pounds \uparrow \$\pounds \downarrow \$\}$ (the same as in Table 9) with no restrictions on the number or type of contractions (the corresponding diagrams are shown in Figure 20). The complete result includes 170 textual diagrams. k is the number of contractions.

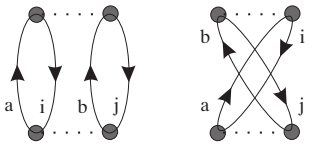
k	Textual diagrams
0	$\square h \square \spadesuit \pounds \downarrow \$\pounds \uparrow \$ \square f \square \heartsuit \pounds \uparrow \$\pounds \uparrow \clubsuit \pounds \downarrow \$\pounds \downarrow \$ \square g \square \heartsuit \pounds \uparrow \$\pounds \uparrow \$\diamondsuit \pounds \uparrow \$\pounds \downarrow \$\diamondsuit \pounds \uparrow \$\pounds \downarrow \$$
1	$\square h \square \spadesuit \pounds \backsim 1\$\pounds \uparrow \$ \square f \square \heartsuit \pounds \uparrow \$\pounds \uparrow \clubsuit \pounds \backsim 1\$\pounds \downarrow \$ \square g \square \heartsuit \pounds \uparrow \$\pounds \uparrow \$\diamondsuit \pounds \uparrow \$\pounds \downarrow \$\diamondsuit \pounds \uparrow \$\pounds \downarrow \$$
2	$\square h \square \spadesuit \pounds \downarrow \$\pounds \backsim 2\$ \square f \square \heartsuit \pounds \backsim 2\$\pounds \uparrow \clubsuit \pounds \backsim 1\$\pounds \downarrow \$ \square g \square \heartsuit \pounds \uparrow \$\pounds \uparrow \$\diamondsuit \pounds \uparrow \$\pounds \backsim 1\$\diamondsuit \pounds \uparrow \$\pounds \downarrow \$$
3	$\square h \square \spadesuit \pounds \downarrow \$\pounds \backsim 2\$ \square f \square \heartsuit \pounds \uparrow \$\pounds \backsim 3\clubsuit \pounds \backsim 1\$\pounds \downarrow \$ \square g \square \heartsuit \pounds \uparrow \$\pounds \uparrow \$\diamondsuit \pounds \backsim 3\$\pounds \downarrow \$\diamondsuit \pounds \backsim 2\$\pounds \backsim 1\$$
4	$\square h \square \spadesuit \pounds \backsim 1\$\pounds \backsim 3\$ \square f \square \heartsuit \pounds \uparrow \$\pounds \backsim 4\clubsuit \pounds \backsim 2\$\pounds \downarrow \$ \square g \square \heartsuit \pounds \uparrow \$\pounds \uparrow \$\diamondsuit \pounds \backsim 4\$\pounds \backsim 2\$\diamondsuit \pounds \backsim 3\$\pounds \backsim 1\$$

6.6 The Translator function

In this section, for simplicity's sake, we will be concerned with scalar expressions only, since it is clear how to extend the shown results to the case of operators. As the batch of textual diagrams is output by MultiContractor, it is probable that some of these diagrams will produce the identical numerical result at arbitrary values of the elements of the corresponding multi-indexed arrays. We will call these diagrams (or algebraic expressions corresponding to them) equivalent. The equivalence may be brought forth by the different symbols for dummy indices (thus, the textual diagrams $\square h \square \clubsuit \pounds \backsim 1\$\pounds \backsim 1\$$ and $\square h \square \clubsuit \pounds \backsim 2\$\pounds \backsim 2\$$ and the corresponding algebraic expressions h_{ii} and h_{jj}, where the summation over the same indices is carried out, are equivalent), by permutational (akin to the well-known symmetries of the two-electron integrals in a physicist's notation $g_{ij,kl} = g_{ji,lk}$) or other symmetries of the operators (for example, $\square h \square \clubsuit \pounds \backsim i\$\pounds \backsim j\$$ is equivalent to $\square h \square \clubsuit \pounds \backsim j\$\pounds \backsim i\$$ if the matrix representing h is symmetric: $h_{ij} = h_{ji}$). As the number of contractions increases, the number of equivalent terms typically grows dramatically and it would be highly desirable to develop a mechanism of their identification. A conventional approach to this problem is to postulate some canonical form for a fully-contracted expression, convert all the terms to this canonical form and simplify the result by means of cancellations, such as $a - a = 0$, and additions, such as $a + a = 2a$.

However, it may not be obvious how to lay down the rules for bringing the expression to a canonical form and in which order to execute them because some of these rules may be non-commutative. Furthermore, the problem is aggravated by the presence in the expression of two or more identical operators (the problem of so-called 'cyclic' contractions), in which case the canonical form may not exist at all, and the canonicalization procedure requires additional operations, some of which scale factorially with the number of dummy indices [150]. *Nostromo*'s Translator function is built on a different approach. It abandons the canonical form altogether and uses the topological structure of the diagrams to decide whether the two given terms are equivalent or not. The transparency and elegance of the diagrammatic simplification of expressions emphasizes the benefits of the use of diagrams in the automatization of second-quantized algorithms. We will demonstrate how our simplification procedure works for scalars (fully-contracted diagrams), but it is straightforward to extend this technique to operators, which may contain terms with uncontracted lines. It may be easily seen why it is advantageous to use a diagrammatic representation, as opposed to an algebraic representation, in tackling the present problem. Firstly, diagrams do not bear dummy indices, which exempts us from the necessity of their reshuffling and renaming. This way, any two expressions that differ by the names of dummy indices only will have the same diagrams. Secondly, by the virtue of the equivalence of all the electrons (or other particles, should the operators be written for them), the location of a particular vertex that belongs to a certain operator is immaterial – it is important only with vertices of which operators, and by means of which lines, this particular vertex is connected. Here we introduce the notion of cycle path as the recording of the journey (the vertices of which operators and by means of which lines (up,down) are encountered) that one makes if one sets off from a particular vertex and finally arrives at it again, forming a cycle. For example, the two equivalent algebraic expressions $g_{ij,ab}g_{ab,ij}$ and $g_{ij,ab}g_{ba,ji}$ that differ by a permutation will have different diagrams (shown in Figure 21), but yet the same cycle paths.

Thirdly, cycles change only their direction but not their structure under the symmetry



(a) (b)

Figure 21: The diagrammatic representation of equivalent algebraic terms $g_{ij,ab}g_{ab,ij}$ (a) and $g_{ij,ab}g_{ba,ji}$ (b).

transformation of the type $h_{i,j} = h_{j,i}$ or $g_{ij,kl} = g_{il,kj}$, and so the backwards-directed cycle paths have to be generated as well, if this type of symmetry is to be taken into account. These observations allow us to conclude that the identity of a fully-contracted mathematical expression is maintained by the number and topological structure of the cycles of its diagrammatic representation. Unfortunately, the present approach does not allow one to handle in the general case special types of symmetry, such as the antisymmetry properties of coupled-cluster amplitudes $t_{ij,ab} = -t_{ij,ba}$, but such special symmetries may be efficiently treated by conventional canonization after the topological analysis has produced the maximally simplified (with respect to the above-mentioned symmetries) expression. The algorithm, from the point when a number of textual diagrams is received from MultiContractor, to the final, maximally simplified algebraic expression, proceeds in the following manner. (a) The first textual diagram is taken and its topological signature (TS) is produced. The TS is an ordered array of the cycle paths of the selected diagram.

Much of the advanced algebraic analysis of the topological structure of diagrams has been done in Ref. [158] (see also references therein) but for our simple purposes we shall not need it in this work. The way of recording cycle paths is to a certain degree optional. We adopted the following convention: the name of the operator from whose vertex (vertex 1) the cycle starts (say, h) is paired up with the index of the line (say, i) that connects it with the subsequent vertex (vertex 2), so that the pair becomes $\{h, i\}$. If the line does not have any index, but is a summation line instead (corresponding to some dummy index in

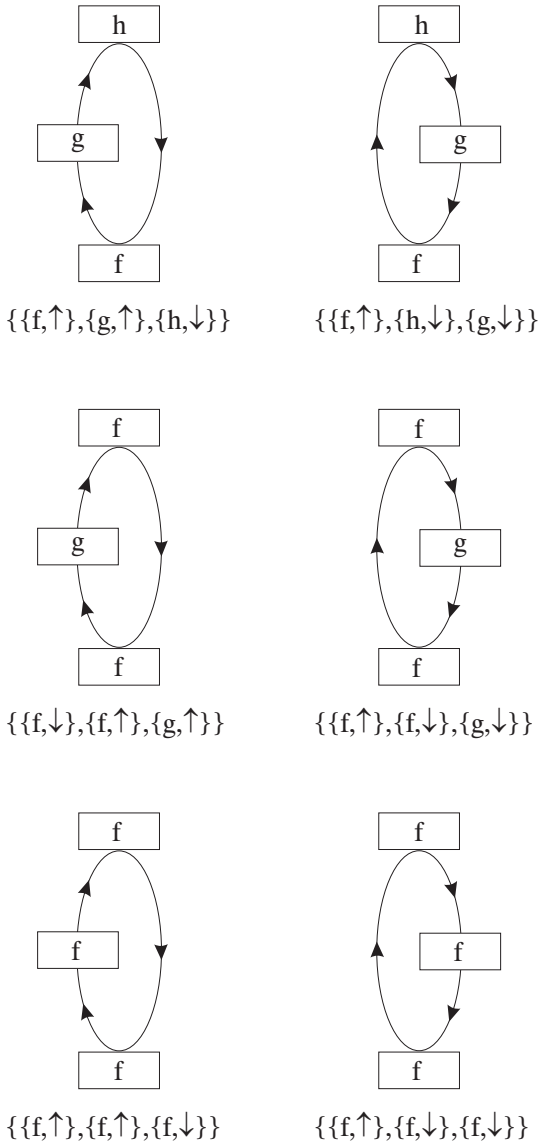


Figure 22: The cycles and their topological signatures. Note that the cycles distinguished by direction only have different topological signatures, and that the topological signatures of the cycles with several identical operators are well-defined.

algebraic notation), its 'index' is written as ↑ or ↓, depending on whether the line faces up or down. The same operations are carried out for the vertex 2 and the line that leaves it, producing the second pair. The process goes on until the cycle ends at vertex 1. Then, all the *cyclic* permutations are generated from the obtained pairings (this operation scales only

as N where N is the length of the cycle) and the permutation that is more ordered than the rest (order may be chosen arbitrarily), is selected as the cycle path. The cyclic permutations are needed to identify equivalent cycles, in case we started the paths from different vertices. Some examples of the cycles and their paths are presented in Figure 22).

When all the cycle paths are thus recorded, they are combined in an ordered way to produce the TS. The TS for the first term is saved along with the index 1 in the TS database; (b) The second textual diagram is taken and its TS [1] is generated. If it is the same as the one already saved, term number 2 is declared equal to term number 1. If the TS of the second term is different from that of the first, it is saved along with number 2 in the TS database. The TS of the third term is compared with all the saved TS's and if does not match any, it is also saved. These operations continue until all the terms are processed, so that the final expression contains the minimal number of identical terms; (c) If any special symmetry within some operators is present, the expression is further simplified by ordering indices that belong to this operator; (d) The sign of the diagram is computed as $(-1)^{N_s+N_h}$ where N_s is the number of cycles (which is of course the number of cycle paths in TS) and N_h is the number of hole lines in the diagram; (e) At this stage only the terms with different TS's exist. It may happen, however, that there will be terms that have identical cycle paths. This means that the group of these terms may be further simplified by factorization, with the common cycle (or several common cycles) factored out; (f) The resulting textual diagrams are converted into traditional algebraic notation, and the final result is produced.

[1] For every cycle in which all of its operators are symmetric, the forward- and backward-directed cycle paths should be generated. Thus, the diagram will be represented by N^2 topological signatures (where N is the number of such cycles) differing by at least one cycle direction. During the comparison with the topological signatures in the TS database, only one of these N^2 topological signatures should match an already saved one for the two diagrams to be declared equivalent. If there is no match, any of the N^2 topological signatures may be saved in the TS database to represent the given diagram.

6.7 Conclusion

In this work we presented a general computer algorithm to manipulate, contract and simplify second-quantized expressions, implemented in the prototype code *Nostromo*. Our approach is topological rather than algebraic and is based on the Goldstone diagrammatic representation of the second-quantized expressions. We showed that using diagrammatic techniques to automate the evaluation of the second-quantized expressions has a number of significant advantages over the traditional approaches which employ strings of creation and annihilation operators and subsequent application of Wick's theorem: firstly, representing diagrams with text strings permits easy and intuitive manipulation of symbolic expressions on a computer; secondly, during the contraction process, the efficient generation of terms with the required structure only (such as fully-contracted terms) is possible; and thirdly, the topological analysis of the diagrams greatly facilitates the simplification of the output results.

The plans for future work include adding flexibility to deriving and manipulating symbolic expressions by extending the functionality of the present code. In particular, we are working on functions that will perform automatic spin integration and visualization of textual diagrams. The presented approach may also be geared towards other diagrams, for example, of Hugenholtz type. *Nostromo* is a first step in the direction of creating a much larger and more functional computer package for automatization of symbolic operations in quantum chemistry.

The material of this chapter was published in the *Journal of Chemical Physics* [159].

CHAPTER VII

SOME SIMPLE RESULTS FOLLOWING FROM LÖWDIN'S PARTITIONING TECHNIQUE

7.1 Introduction

Throughout this chapter, all our equations are written in a finite-dimensional, matrix-vector form. Löwdin's partitioning technique [160, 161, 162, 3] which we briefly outline below, is an important tool in the development of quantum chemical formalisms and is actively applied in contemporary research [163, 56, 164, 165, 166, 167]. The variation of the Rayleigh-Ritz functional

$$E = \langle C|H|C\rangle / \langle C|C\rangle \tag{110}$$

written in terms of the discrete variational configuration interaction coefficients C_i results in the well-known eigenvalue equation which for the purpose of partitioning is written in the following block matrix form

$$\left(\begin{array}{c:c} H^{QQ} & H^{QR} \\ \hdashline H^{RQ} & H^{RR} \end{array}\right) \left(\begin{array}{c} C^Q \\ \hdashline C^R \end{array}\right) = E \left(\begin{array}{c} C^Q \\ \hdashline C^R \end{array}\right), \tag{111}$$

or, multiplying the blocks,

$$H^{QQ}|C^Q\rangle + H^{QR}|C^R\rangle = E|C^Q\rangle, \tag{112}$$

$$H^{RQ}|C^Q\rangle + H^{RR}|C^R\rangle = E|C^R\rangle. \tag{113}$$

The superscripts in these equations have the meaning of the dimension of the corresponding quantities. The total dimension of the system (111) is $N = Q + R$. To arrive at the effective eigenvalue problem, which is the central equation of the partitioning technique, we use the second of these equations to express C^R through C^Q

$$|C^R\rangle = (E - H^{RR})^{-1} H^{RQ}|C^Q\rangle, \tag{114}$$

we eliminate C^R and arrive at the representation

$$\left[H^{QQ} + H^{QR}(E - H^{RR})^{-1}H^{RQ}\right]|C^Q\rangle = E|C^Q\rangle \tag{115}$$

which is equivalent to (111). The partitioning technique may be used for a direct solution of the eigenvalue problem. Various expansion techniques [160] allow us to write the energy E contained in this non-linear equation as an infinite series of terms. It is possible to derive in such a way, for example, the Brillouin-Wigner and Rayleigh-Schrödinger perturbation formulas. When E and C^Q have been determined from (115), the remaining component C^R is calculated from (114). Assuming that $(E - H^{RR})^{-1}$ exists and the perturbation series used to calculate E converges, the equations (111) are solved. Although in practice the perturbation series is terminated at the second or third order and E and C are determined with an error, it is important to observe that this approach in principle provides the exact solution of (111).

7.2 Discussion

The usefulness of the partitioning technique rests on its ability to find C and E which are the solutions of the eigenvalue problem (111). If, however, we assume that C and E are already known, a few interesting mathematical results follow from equations (111). In what follows we prove some results concerning the spectral properties of matrices with eigenvectors whose parts are linearly dependent.

Surprisingly, it is possible to write the formal solution of the eigenvalue equations in a closed form. Multiplying (113) by H^{QR} from the left (so that the matrix acting on C^Q becomes square) and then expressing C^Q through E as in

$$|C^Q\rangle = W^{QR}\left(E - H^{RR}\right)|C^R\rangle, \tag{116}$$

$$W^{QR} = (H^{QR}H^{RQ})^{-1}H^{QR}, \tag{117}$$

we rewrite (112) as a quadratic equation in E:

$$E^2|a\rangle + E|b\rangle + |c\rangle = 0, \tag{118}$$

where the vectors $|a\rangle$, $|b\rangle$, and $|c\rangle$ are defined as

$$|a\rangle = -W^{QR}|C^R\rangle, \tag{119}$$

$$|b\rangle = H^{QQ}W^{QR}|C^R\rangle - W^{QR}|U^R\rangle, \tag{120}$$

$$|c\rangle = H^{QQ}W^{QR}|U^R\rangle - |U^Q\rangle, \tag{121}$$

and the following abbreviations are used

$$|U^Q\rangle = -H^{QR}|C^R\rangle \tag{122}$$

$$|U^R\rangle = -H^{RR}|C^R\rangle. \tag{123}$$

The assumption that the inverse of $H^{QR}H^{RQ}$ exists leads us to the necessary condition $Q \leq R$ (which will be adhered to in the rest of the chapter, unless indicated otherwise). For convenience of discussion, the Q equations which constitute (118) may also be 'averaged' by means of a contraction with an arbitrary vector $\langle X^Q|$:

$$\langle X^Q|a\rangle E^2 + \langle X^Q|b\rangle E + \langle X^Q|c\rangle = 0. \tag{124}$$

Clearly, when C^R is exact, the root E of the projected equation (124), which is at the same time the root of (111), does not depend on the particular choice of $\langle X^Q|$. Note that for a two-by-two matrix H equation (118) is a scalar one and is identical to the characteristic equation of H.

Suppose we choose such an eigenvector C_i of H that has a unique C^R. In other words, no other eigenvector of H shares the part C^R with our chosen eigenvector. The eigenvalue corresponding to this eigenvector is E_i. Then the projected equation (124) constructed from this C^R will have two roots, one of which is guaranteed to be E_i. Another root does not have to satisfy the eigenvalue equations and in the general case depends on the projection

$\langle X^Q|$. The independence of projection may allow us to distinguish the true solution of (111) from the spurious one. The spurious root has its origin in the act of projection by $\langle X^Q|$ and its existence is explained by the fact that the scalar equations which constitute (118) are satisfied by the actual eigenvalue of (111) only. Assume now that we have two eigenvectors C_i and C_j with two distinct eigenvalues E_i and E_j and two identical parts C^R. In this case the individual scalar equations in (118) may be satisfied by two different roots, and the projected equation (124) will have two solutions, E_i and E_j which do not depend on the projection.

Now we would like to make a comment that will be used in the proof of the results below. It is well-known that two eigenvectors $|d_1\rangle$ and $|d_2\rangle$ of $N\mathrm{x}N$ matrix A that correspond to the same eigenvalue E may be added together with arbitrary coefficients α_1 and α_2, and the vector $|d_3\rangle$ resulting from this linear combination will also be an eigenvector of A with the same eigenvalue E. Suppose that we want to take two linear combinations of $|d_1\rangle$ and $|d_2\rangle$ in such a way that two resulting eigenvectors $|d'_1\rangle$ and $|d'_2\rangle$ are partly coincident:

$$|d'_1\rangle - |d'_2\rangle = \{d'_{11} - d'_{21}, d'_{12} - d'_{22}, ..., d'_{1k} - d'_{2k}, 0, 0, ..., 0, d'_{1l} - d'_{2l}, d'_{1l+1} - d'_{2l+1}, ..., d'_{1N} - d'_{2N}\}. \tag{125}$$

In this formula, d'_{1k} is the k-th component of the vector $|d'_1\rangle$. Obviously, this is possible only when $l-k-1$ corresponding components of $|d_1\rangle$ and $|d_2\rangle$ constitute linearly dependent vectors, i.e. if at least one of these vectors is zero or

$$\{d_{1,k+1}, d_{1,k+2}, ..., d_{1,l-1}\} = \alpha\{d_{2,k+1}, d_{2,k+2}, ..., d_{2,l-1}\}, \tag{126}$$

where $\alpha \neq 0$. If the $l-k-1$ components of $|d_1\rangle$ and $|d_2\rangle$ constitute linearly independent vectors, two different vectors $|d'_1\rangle$ and $|d'_2\rangle$ which satisfy (125) cannot be constructed because a vector has unique coordinates in the basis of $|d_1\rangle$ and $|d_2\rangle$. So, it appears that if two eigenvectors $|d_1\rangle$ and $|d_2\rangle$ have linearly dependent parts, we may always convert them into the eigenvectors $|d'_1\rangle$ and $|d'_2\rangle$ which have the coincident parts, so that $|d'_1\rangle$ and $|d'_2\rangle$ remain eigenvectors with the eigenvalues of $|d_1\rangle$ and $|d_2\rangle$, respectively.

Theorem 1. If an arbitrary square matrix H separated into blocks in the manner of that in equation (111) (so that $R \geq Q$) possesses two eigenvectors C_1, C_2 with the same eigenvalue and with linearly dependent parts C_1^R, C_2^R, then $\mathrm{rank}(H^{QR}) < Q$.

Proof. Suppose $\mathrm{rank}(H^{QR}) = Q$, then $H^{QR}H^{RQ}$ is invertible and expression (116) can be constructed. Further, C_1 and C_2 may be converted into the eigenvectors C_1' and C_2' by two suitable linear combinations so that $C_1'^R = C_2'^R \equiv C^R$. Then (by equation 116) C^Q is a linear, deterministic function of E and C^R. This contradicts the fact that different vectors C^Q may correspond to one eigenvalue E and one vector C^R. Thus, our supposition is not true and $\mathrm{rank}(H^{QR}) < Q$.

As an important corollary used in the proof of the next theorem, we note that if an arbitrary square matrix H separated into blocks in the manner of that in equation (111) (so that $R \geq Q$) possesses two eigenvectors C_1 and C_2 with linearly dependent parts C_1^R, C_2^R and $\mathrm{rank}(H^{QR}) = Q$, then two eigenvalues corresponding to C_1 and C_2 are non-degenerate. Indeed, under the conditions of Theorem 1, if the two eigenvalues are degenerate, then $\mathrm{rank}(H^{QR}) < Q$. If they are not degenerate, $\mathrm{rank}(H^{QR})$ may be either Q or less than Q. So, we conclude that the equality $\mathrm{rank}(H^{QR}) = Q$ may correspond only to the situation in which the two eigenvalues are non-degenerate.

Now we prove another theorem related to the possibility of having three eigenvectors with identical parts C^R which further explains some properties of equation (118).

Theorem 2. If an arbitrary square matrix H separated into blocks in the manner of that in equation (111) (so that $R \geq Q$) possesses three eigenvectors C_1, C_2, C_3 with linearly dependent parts C^R, then $\mathrm{rank}(H^{QR}) < Q$.

Proof. Suppose $\mathrm{rank}(H^{QR}) = Q$. Then $H^{QR}H^{RQ}$ is invertible and (118-124) can be constructed. By theorem 1, three eigenvalues corresponding to the three eigenvectors C_1, C_2, C_3, are distinct. By suitable multiplications we may convert C_1, C_2, C_3 into the eigenvectors C_1', C_2', C_3' with $C_1'^R = C_2'^R = C_3'^R \equiv C^R$. The projected quadratic equation (124) is satisfied with all possible eigenvalues of H that correspond to C_1, C_2, C_3. However,

since the equation is quadratic, it cannot be satisfied with three distinct values of E. Hence, we arrive at the contradiction and $\mathrm{rank}(H^{QR}) < Q$.

Another simple conclusion follows directly from equation (114).

Theorem 3. If an arbitrary square matrix H separated into blocks in the manner of that in equation (111) possesses two eigenvectors C_1 and C_2 with linearly dependent parts C_1^Q and C_2^Q (Q is arbitrary) and identical eigenvalues E, then E is also an eigenvalue of the matrix H^{RR}.

Proof. Suppose $(E - H^{RR})^{-1}$ exists. The vectors C_1^Q and C_2^Q may be equated ($C_1^Q = C_2^Q \equiv C^Q$) by two suitable linear combinations. Then C^R is a linear, deterministic function of E and C^Q. This contradicts the fact that C_1^R, C_2^R correspond to only one vector C^Q and one value E. The only way to resolve this contradiction is to assume that (114) cannot be constructed because $(E - H^{RR})^{-1}$ does not exist. Consequently, E is the eigenvalue of H^{RR}.

Note that using this theorem for the case when $Q = 1$ we arrive at the corollary which also follows from the separation theorem by MacDonald [168]: if the matrix $||H_{ij}||_{1 \le i,j \le N}$ possesses degenerate eigenvalue E the matrix $||H_{ij}||_{1 \le i,j \le N-1}$ possesses this eigenvalue too.

The material of this chapter is in press in the *Journal of Mathematical Chemistry* [169].

CHAPTER VIII

CONCLUSION

The central work of this thesis presented in chapters 3-4 features the investigation of hybrid quantum-chemical methods, in which an inexpensive but not very accurate method is fused with a more resource-demanding but more accurate approach.

In the beginning we examine which terms in the MP2 energy expression (37) contribute most to its divergent behavior at large interatomic separations. For this purpose we divide the orbital space into four different non-overlapping sets: restricted occupied, active occupied, active virtual and restricted virtual (in most cases the active space spans just the valence orbitals). Then we explore nine different energy contributions resulting from double excitations from the occupied onto the virtual orbitals as a function of interatomic distance. The investigated systems are small molecules (often diatomics) for which the full configuration calculation treatment is possible in a reasonable basis set 6-31G*. This allows us to compare the results of our method to the exact results (of course, within the limit of the 6-31G* basis set). As expected, and confirmed in our studies, the most profound effect on the divergence is due to the few excitations from the active occupied orbitals onto the active virtual orbitals. The rest of the excitations, when compared to their CCSD analogs, show a qualitatively similar behavior, and if not, are of little numerical importance.

Judging from this picture, we decided to remove these 'noxious' energy contributions from the MP2 energy expression and instead use the corresponding CCSD terms. If the hybrid scheme, fusing MP2 and CCSD, is to be of any help, it has to combine the inexpensiveness of MP2 and the accuracy of CCSD. Because of the first requirement, we are not allowed to perform an expensive CCSD calculation in order to generate the CCSD contributions arising completely from the active space excitations. We can only afford to solve

the CCSD equations in the active space since the dimension of the active space is small and this computation will not add much to the cost of the initial MP2 expenditure. Our experiments clearly demonstrated that the 'mechanical' combination of MP2 and active-space CCSD energy contributions, when the two methods are not 'aware' of each other, results in energy curves which still diverge albeit not in such a pronounced manner than those of MP2. However, if we solve the for the CCSD amplitudes in active space but *in the presence* of the MP2 amplitudes, the resulting energy curves are physically correct. Indeed, they are almost parallel to the CCSD energy curves.

An extra improvement of the described hybrid scheme comes from the realization that describing some other energy amplitudes at the coupled cluster level will not affect the overall $\mathcal{O}(N^5)$ scaling of the method. The dissociation curves produced by this improved hybrid method are essentially parallel to those produced by CCSD. This is certainly a very desirable performance for a $\mathcal{O}(N^5)$ method, since the analogously scaling popular methods (MP2, Epstein-Nesbet second order perturbation theory and CC2) fail at the dissociation limit.

We also need to attract attention to the limitations of our hybrid models. The most obvious drawback is the need to choose the active space before any of our hybrid methods can be used. We assume (and prove) that for many cases the minimal or very small active space will suffice, but the inconvenience of choosing the adequate active space remains, since it does not always coincide with or even include HOMO and LUMO. One possible way of automating the process of selecting the chemically-important excitations that will be treated with higher accuracy is to select them based on the analysis of the occupation numbers. Using these numbers, we can identify the small set of orbitals (and, consecutively, excitations) which are most important for the description of the dissociation process. This approach is not without problems either. First, how to choose the optimal cut-off value for the occupation numbers? And second, since the choice of the excitations will depend on the geometry, will this inconsistency preserve the smoothness of potential energy curves (a

similar difficulty emerges in local correlation theories [170]).

One more possible limitation of the proposed hybrid methods is that all of them are expected to work in cases where the CCSD method works. Where CCSD fails, the hybrid methods based on it may be helpless. For example, the proper dissociation of the N_2 molecule may require up to the CCSDTQPH (coupled cluster with up to sixtuple excitations in (35)) treatment [82]. Certainly, one can hope to successfully combine MP2 with better-quality coupled-cluster methods. We made some initial experiments in combining MP2 with CCSDTQ for the symmetric dissociation of water, but we need a more systematic study before we can make a conclusion about the performance or usefulness of such 'advanced' hybrid methods.

An exciting prospect for the hybrid methods of the kind developed in this work is their application for the non-covalent interactions, such as dispersion interactions. Large molecular systems, for example, benzene dimer, are very resource-demanding, so that the MP2 calculations are often feasible but the CCSD and especially CCSD(T) calculations are prohibitively expensive [171, 172, 173]. The accurate potential energy curves for such interactions are required to generate the interaction potentials which are needed in such areas as drug discovery and dynamical simulations. For dispersion interactions, no covalent bond is broken when one fragment recedes from another, and no degeneracy in orbitals occurs. Therefore even MP2 method, which fails for the dissociation of covalent bonds, works for the dispersion interactions. Nevertheless, it would be very desirable to generate the curves for non-covalent interaction with higher accuracy at essentially the same computational cost. The hybrid methods might need a refurbishing for this task: the choice of the active space will be based either on the occupation numbers analysis or the method itself will utilize the natural orbitals instead of Hartree-Fock ones. Another possible application of the hybrid methods is mentioned in the conclusion to Chapter 4.

Chapters 5-7 of the thesis describe projects which are not related to the development

of hybrid methods. They concern diverse topics: non-adiabatic quantum chemistry, diagrammatic technique, and Löwdin's partitioning technique. These works are nevertheless included here since they were performed at the time of my graduate studentship and give useful summary of the knowledge and skills which I have acquired.

REFERENCES

[1] Cook, D. B., *Handbook of Computational Quantum Chemistry*, Oxford University Press, New York, 1998.

[2] Quinn, M., *Computational Quantum Chemistry: An Introduction to Basis Set Theory*, Academic Press, San Diego, 2001.

[3] McWeeny, R., *Methods of Molecular Quantum Mechanics*, Academic Press, London, 1989.

[4] Szabo, A. and Ostlund, N. S., *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, McGraw-Hill, New York, 1989.

[5] Jensen, F., *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, John Wiley & Sons, Chichester, 1999.

[6] Fock, V., Z. Physik **61** (1930) 126.

[7] Fock, V., Z. Physik **62** (1930) 795.

[8] Hartree, D. R., Proc. Cambridge Philos. Soc. **45** (1948) 230.

[9] Roothaan, C. C. J., Rev. Mod. Phys. **23** (1951) 69.

[10] Roothaan, C. C. J., Rev. Mod. Phys. **32** (1960) 179.

[11] Shavitt, I., The method of configuration interaction, in *Methods of Electronic Structure Theory*, page 189, New York, 1977, Plenum Press.

[12] Sherrill, C. D. and Schaefer, H. F., Adv. Quant. Chem. **34** (1999) 143.

[13] Coester, F., Nucl. Phys. **1** (1958) 421.

[14] Coester, F. and Kümmel, H., Nucl. Phys. **17** (1960) 477.

[15] Čížek, J., J. Chem. Phys. **45** (1966) 4256.

[16] Čížek, J., Adv. Chem. Phys. **14** (1969) 35.

[17] Čížek, J. and Paldus, J., Int. J. Quant. Chem. **5** (1971) 359.

[18] Bartlett, R. J., J. Phys. Chem. **93** (1989) 1697.

[19] Monkhorst, H., Int. J. Quant. Chem. Sympos. **11** (1977) 421.

[20] Mukherjee, D. and Mukherjee, P. K., Chem. Phys. **39** (1979) 325.

[21] Emrich, K., Nucl. Phys. A **351** (1981) 379.

[22] Ghosh, S. and Mukherjee, D., Proc. Ind. Acad. Sci. **93** (1984) 947.

[23] Sekino, H. and Bartlett, R. J., Int. J. Quant. Chem. Symp. **18** (1984) 255.

[24] Geertsen, J., Rittby, M., and Bartlett, R. J., Chem. Phys. Lett. **164** (1989) 57.

[25] Brillouin, L., J. Phys. Radium **3** (1932) 373.

[26] Møller, C. and Plesset, M. S., Phys. Rev. **46** (1934) 618.

[27] Löwdin, P.-O., J. Chem. Phys. **19** (1951) 1396.

[28] Bartlett, R. J. and Silver, D. M., J. Chem. Phys. **62** (1975) 3258.

[29] Born, M. and Oppenheimer, R., Ann. Phys. (Paris) **84** (1927) 457.

[30] Weyl, H., *The Classical Groups, Their Invariants and Representations*, Princeton University Press, Princeton, 1946.

[31] Adams, W. H., Phys. Rev. **127** (1962) 1650.

[32] Thouless, D. J., Nucl. Phys. **21** (1960) 225.

[33] Wilson, S., *Electron Correlation in Molecules*, Clarendon, Oxford, 1984.

[34] Harris, F. E., Monkhorst, H. J., and Freeman, D. L., *Algebraic and Diagrammatic Methods in Many-Fermion Theory*, Oxford University Press, New York, 1992.

[35] Zaitsevsky, A. V., *Methods in Many-Body Theory in Quantum Chemistry*, Moscow University Press, Moscow, 1993, in Russian.

[36] Lindgren, I. and Morrison, J., *Atomic Many-Body Theory*, Springer, New York, 1982.

[37] Jørgensen, P. and Simons, J., *Second Quantization-Based Methods in Quantum Chemistry*, Academic Press, New York, 1981.

[38] Goldstone, J., Proc. Roy. Soc. (London) **239** (1957) 267.

[39] Löwdin, P. O., Phys. Rev. **97** (1955) 1509.

[40] Bartlett, R. J. and Stanton, J. F., Applications of post-hartree-fock methods: A tutorial, in *Reviews in Computational Chemistry*, volume 5, page 65, New York, 1977, VCH Publishers.

[41] Dutta, A. and Sherrill, C. D., J. Chem. Phys. **118** (2003) 1610.

[42] Sherrill, C. D., Annu. Rep. Comp. Chem. **1** (2004) 45.

[43] Abrams, M. L. and Sherrill, C. D., J. Chem. Phys. **121** (2004) 9211.

[44] Abrams, M. L. and Sherrill, C. D., Chem. Phys. Lett. **404** (2005) 284.

[45] J. Olsen, P. Jørgensen, H. K. A. B., J. Chem. Phys. **104** (1996) 8007.

[46] Larsen, H., Olsen, J., Jørgensen, P., and Christiansen, O., J. Chem. Phys. **113** (2000) 6677.

[47] Larsen, H., Olsen, J., Jørgensen, P., and Christiansen, O., J. Chem. Phys. **114** (2001) 10985(E).

[48] Krylov, A. I., Chem. Phys. Lett. **338** (2001) 375.

[49] Krylov, A. I. and Sherrill, C. D., J. Chem. Phys. **116** (2002) 3194.

[50] Shao, Y. H., Head-Gordon, M., and Krylov, A. I., J. Chem. Phys. **118** (2003) 4807.

[51] Sears, J. S., Sherrill, C. D., and Krylov, A. I., J. Chem. Phys. **118** (2003) 9084.

[52] Piecuch, P., Kowalski, K., Pimienta, I. S. O., and Kucharski, S. A., Method of moments coupled-cluster equations: A new theoretical framework for designing "black box" approaches for molecular potential energy surfaces, in *Low-Lying Potential Energy Surfaces*, ACS Symposium Series, volume 828, page 31, Washington, D.C., 2002, American Chemical Society.

[53] P. Piecuch, K. Kowalski, I. S. O. P. and McGuire, M. J., Int. Rev. Phys. Chem. **21** (2002) 527.

[54] Crawford, T. D. and Stanton, J., Int. J. Quant. Chem. **70** (1998) 601.

[55] Li, X. and Paldus, J., J. Chem. Phys. **108** (1998) 637.

[56] Gwaltney, S. R., Sherrill, C. D., Head-Gordon, M., and Krylov, A. I., J. Chem. Phys. **113** (2000) 3548.

[57] Gwaltney, S. R. and Head-Gordon, M., J. Chem. Phys. **115** (2001) 2014.

[58] Tobita, M. et al., J. Chem. Phys. **119** (2003) 10713.

[59] Saebø, S. and Pulay, P., J. Chem. Phys. **88** (1988) 1884.

[60] Hampel, C. and Werner, H.-J., J. Chem. Phys. **104** (1996) 6286.

[61] Crawford, T. D. and King, R. A., Chem. Phys. Lett. **366** (2002) 611.

[62] Purvis, G. D. and Bartlett, R. J., Chem. Phys. Lett. **76** (1982) 1910.

[63] Crawford, T. D. et al., J. Comp. Chem. submitted.

[64] Nooijen, M., J. Chem. Phys. **111** (1999) 10815.

[65] Olsen, J., Roos, B. O., Jørgensen, P., and Jensen, H. J., J. Chem. Phys. **89** (1988) 2185.

[66] Piecuch, P., Oliphant, N., and Adamowicz, L., J. Chem. Phys. **99** (1993) 1875.

[67] Sherrill, C. D. and Schaefer, H. F., J. Phys. Chem. **100** (1996) 6069.

[68] Klopper, W., Noga, J., Koch, H., and T.Helgaker, Theor. Chem. Acc. **97** (1997) 164.

[69] P. Piecuch, S. A. K. and Bartlett, R. J., J. Chem. Phys. **110** (1999) 6103.

[70] Li, X. and Paldus, J., Mol. Phys. **98** (2000) 1185.

[71] Crawford, T. D., Schaefer, H. F., and Lee, T. J., J. Chem. Phys. **105** (1996) 1060.

[72] Stanton, J. F., Gauss, J., Watts, J. D., and Bartlett, R. J., Mol. Phys. **98** (2000) 1185.

[73] Bochevarov, A. D. and Sherrill, C. D., J. Chem. Phys. **122** (2005) 234110.

[74] Moscardó, F., Pérez-Jiménez, A. J., Sancho-García, and San Fabián, E., Chem. Phys. Lett. **288** (1998) 418.

[75] Voorhis, T. V. and Head-Gordon, M., Chem. Phys. Lett. **317** (2000) 575.

[76] Voorhis, T. V. and Head-Gordon, M., J. Chem. Phys. **112** (2000) 5633.

[77] Beran, G. J. O., Head-Gordon, M., and Gwaltney, S. R., J. Chem. Phys. **124** (2006) 114107.

[78] Krylov, A. I., Chem. Phys. Lett. **350** (2001) 522.

[79] Kowalski, K. and Piecuch, P., J. Chem. Phys. **116** (2002) 7411.

[80] Włoch, M., Gour, J. R., Kowalski, K., and Piecuch, P., J. Chem. Phys. **122** (2005) 214107.

[81] Raghavachari, K., Trucks, G. W., Pople, J. A., and Head-Gordon, M., Chem. Phys. Lett. **157** (1989) 479.

[82] Musial, M. and Bartlett, R. J., J. Chem. Phys. **122** (2005) 224102.

[83] Christiansen, O., Koch, H., and Jørgensen, P., Chem. Phys. Lett. **243** (1995) 409.

[84] Larsen, H., Olsen, J., Jørgensen, P., and Gauss, J., Chem. Phys. Lett. **342** (2001) 200.

[85] Hättig, C., J. Chem. Phys. **118** (2003) 7751.

[86] Chen, W. and Schlegel, W. B., J. Chem. Phys. **101** (1994) 5957.

[87] Krylov, A. I., J. Chem. Phys. **113** (2000) 6052.

[88] Meyer, W., J. Chem. Phys. **58** (1973) 1017.

[89] Kutzelnigg, W., Pair correlation theories, in *Methods of Electronic Structure Theory*, volume 828, page 129, New York, 1977, Plenum Press.

[90] Murray, C. and Davidson, E. R., Int. J. Quant. Chem. **43** (1992) 755.

[91] Temelso, B., Sherrill, C. D., Merkle, R. C., and Freitas, J. R. A., J. Phys. Chem. A submitted.

[92] Levine, I. N., *Molecular Spectroscopy*, Wiley & Sons, New York, 1975.

[93] Temelso, B., Valeev, E. F., and Sherrill, C. D., J. Phys. Chem. A **108** (2004) 3068.

[94] Bochevarov, A. D., Temelso, B., and Sherrill, C. D., J. Chem. Phys. submitted.

[95] Cafiero, M. and Adamowicz, L., J. Chem. Phys. **116** (2002) 5557.

[96] Cafiero, M. and Adamowicz, L., Phys. Rev. Lett. **89** (2002) 073001.

[97] Bubin, S. and Adamowicz, L., J. Chem. Phys. **118** (2003) 3079.

[98] Cafiero, M., Bubin, S., and Adamowicz, L., Phys. Chem. Chem. Phys. **5** (2003) 1491.

[99] Wolniewicz, L., Chem. Phys. Lett. **24** (2003) 461.

[100] Wolniewicz, L., J. Chem. Phys. **103** (1995) 1792.

[101] Handy, N. C., Yamaguchi, Y., and Schaefer, H. F., J. Chem. Phys. **84** (1986) 4481.

[102] Handy, N. C. and Lee, A. M., Chem. Phys. Lett. **252** (1996) 425.

[103] Schwenke, D. W., J. Chem. Phys. **114** (2003) 1693.

[104] Schwenke, D. W., J. Phys. Chem. A. **114** (2003) 2352.

[105] Valeev, E. F. and Sherrill, C. D., JCP **118** (2003) 3921.

[106] Deumens, E., Diz, A., Longo, R., and Öhrn, Y., Rev. Mod. Phys. **66** (1994) 917.

[107] Deumens, E. and Öhrn, Y., J. Phys. Chem. A **105** (2001) 2660.

[108] Thomas, I. L., Phys. Rev. **90** (1969) 185.

[109] Thomas, I. L. and Joy, H. W., Phys. Rev. A **2** (1970) 1200.

[110] Tachikawa, M., Mori, K., Nakai, H., and Iguchi, K., Chem. Phys. Lett. **290** (1998) 437.

[111] Tachikawa, M. and Osamura, Y., Theor. Chem. Acc. **104** (2000) 29.

[112] Tachikawa, M., Mori, K., and Osamura, Y., Mol. Phys. **96** (1999) 1207.

[113] Ishida, M., Tachikawa, M., Tokiwa, H., Mori, K., and Ishii, A., Surf. Sci. **438** (1999) 47.

[114] Tachikawa, M., Mol. Phys. **100** (2002) 881.

[115] Nakai, H., Int. J. Quant. Chem. **86** (2002) 511.

[116] Nakai, H., Sodeyama, K., and Hoshino, M., Chem. Phys. Lett. **345** (2001) 118.

[117] Webb, W. P., Iordanov, T., and Hammes-Schiffer, S., J. Chem. Phys. **117** (2001) 4106.

[118] Tachikawa, M., Chem. Phys. Lett. **360** (2002) 494.

[119] Nakai, H. and Sodeyama, K., J. Chem. Phys. **118** (2003) 1119.

[120] Monkhorst, H. J., Int. J. Quant. Chem. **72** (1999) 281.

[121] Kato, T., Commun. Pure. Appl. Math. **10** (1957) 151.

[122] Sutcliffe, B. T., J. Mol. Struct. (THEOCHEM) **341** (1995) 217.

[123] Kutzelnigg, W., Mol. Phys. **90** (1997) 909.

[124] Woolley, R. G. and Sutcliffe, B. T., Chem. Phys. Lett. **45** (1976) 393.

[125] Woolley, R. G. and Sutcliffe, B. T., Chem. Phys. Lett. **73** (1980) 393.

[126] Pulay, P., J. Comp. Chem. **3** (1982) 556.

[127] Wilson, S., in *Ab Initio Methods in Quantum Chemistry – I*, page 439, New York, 1987, John Wiley & Sons.

[128] Bishop, D. and Cheung, L. M., Adv. Quant. Chem. **12** (1980) 1.

[129] Kozlowski, P. M. and Adamowicz, L., Chem. Rev. **93** (1993) 2007.

[130] Zheng, D. C., Barrett, B. R., Jaqua, L., Vary, J. P., and McCarthy, R., Phys. Rev. C **48** (1993) 1083.

[131] Dean, D. J. et al., Phys. Rev. C **59** (1999) 2174.

[132] Sinnokrot, M. O. and Sherrill, C. D., J. Chem. Phys. **115** (2001) 2439.

[133] Martin, J. M. L., Chem. Phys. Lett. **292** (1998) 411.

[134] Martin, J. M. L., Chem. Phys. Lett. **283** (1998) 283.

[135] Thomas, I. L., Phys. Rev. A **3** (1971) 565.

[136] Thomas, I. L., J. Chem. Phys. **118** (2003) 9489.

[137] Bochevarov, A. D., Valeev, E. F., and Sherrill, C. D., Mol. Phys. **102** (2004) 111.

[138] Janssen, C. L. and H. F. Schaefer, III.

[139] Kucharski, S. A. and Bartlett, R. J., Theor. Chim. Acta. **80** (1991) 387.

[140] Li, X. and Paldus, J., J. Chem. Phys. **101** (1994) 8812.

[141] Gotoh, M., Mori, K., and Itoh, R., Int. J. Quant. Chem. **56** (1995) 163.

[142] Crawford, T. D., Lee, T. J., and H. F. Schaefer III, J. Chem. Phys. **107** (1997) 7943.

[143] Crawford, T. D., *Ph.D. thesis*, University of Georgia, 1996.

[144] Harris, F. I., Int. J. Quant. Chem. **75** (1999) 593.

[145] Hirata, S. and Bartlett, R. J., Chem. Phys. Lett. **321** (2000) 216.

[146] Nooijen, M. and Lotrich, V., J. Chem. Phys. **113** (2000) 494.

[147] Nooijen, M. and Lotrich, V., J. Chem. Phys. **113** (2000) 4549.

[148] Nooijen, M. and Lotrich, V., J. Mol. Struct. (THEOCHEM) **547** (2001) 253.

[149] Hirata, S., *Tensor Contraction Engine: A Symbolic Quantum-Mechanical Many-Electron Theories, version 1.0*, Pacific Northwest National Laboratory: Richland, WA, 2002.

[150] Hirata, S., J. Phys. Chem. A **107** (2003) 9887.

[151] Hugenholtz, N. M., Physica **23** (1957) 481.

[152] Kutzelnigg, W., J. Chem. Phys. **82** (1985) 4166.

[153] Crawford, T. D. and H. F. Schaefer III, Pair correlation theories, in *Methods of Electronic Structure Theory*, volume 14, New York, 2000, John Wiley & Sons.

[154] Champaign, IL, *Wolfram Research, Inc., Mathematica, Version 5.0*, 2003.

[155] Paldus, J. and Wong, H. C., Comp. Phys. Commun. **6** (1973) 1.

[156] Paldus, J. and Wong, H. C., Comp. Phys. Commun. **6** (1973) 9.

[157] Kaldor, U., J. Comp. Phys. **20** (1976) 432.

[158] Kvasnička, V., Int. J. Quant. Chem. **21** (1982) 1003.

[159] Bochevarov, A. D. and Sherrill, C. D., J. Chem. Phys. **121** (2004) 3374.

[160] Löwdin, P.-O., J. Chem. Phys. **19** (1951) 1396.

[161] Löwdin, P.-O., J. Math. Phys **3** (1962) 969.

[162] Löwdin, P.-O., J. Mol. Spectr **10** (1963) 12.

[163] Staroverov, V. N. and Davidson, E. R., Chem. Phys. Lett. **43** (1998) 1586.

[164] Surján, P. R. and Szabados, A., Int. J. Quant. Chem. **90** (2002) 20.

[165] Meißner, H., J. Chem. Phys. **119** (2003) 4126.

[166] Lograd o, P. G. and Vianna, J. D. M., J. Math. Chem. **22** (1997) 107.

[167] Maniero, A. M., Rocha Neto, J. F., Malbouisson, L. A. C., and Vianna, J. D. M., Int. J. Quant. Chem. **90** (2002) 1586.

[168] MacDonald, J. K. L., Phys. Rev. **43** (1933) 830.

[169] Bochevarov, A. D. and Sherrill, C. D., J. Math. Chem. (2006) in press.

[170] Russ, N. J. and Crawford, T. D., J. Chem. Phys. **121** (2004) 691.

[171] Sinnokrot, M. O., Valeev, E. F., and Sherrill, C. D., J. Am. Chem. Soc. **124** (2002) 10887.

[172] Sinnokrot, M. O. and Sherrill, C. D., J. Am. Chem. Soc. **126** (2004) 7690.

[173] Sinnokrot, M. O. and Sherrill, C. D., J. Phys. Chem. A **108** (2004) 10200.